

G

Département de géomatique appliquée
Faculté des lettres et sciences humaines
Université de Sherbrooke

G
70.45
US
C66
2012

**La détection des cyanobactéries en milieu lacustre
par l'étude des anomalies des spectres de réflectance de l'eau.**

Gabriel Constantin 1985 -

**Mémoire présenté pour l'obtention du grade de Maître en sciences géographiques (M.Sc.),
cheminement en Télédétection**

Juillet 2012

© Gabriel Constantin, 2012

Directeur de recherche : Yannick Huot

Affiliation : Département de Géomatique Appliquée et Centre d'applications et de recherches en télédétection (Cartel)

Codirecteur de recherche : Norm O'Neill

Affiliation : Département de Géomatique Appliquée, Cartel et Canadian Network for the Detection of Atmospheric Change (CANDAC)

Membres du jury :

Norm O'Neill

Yannick Huot

Alain Royer (Département de Géomatique Appliquée et Cartel, Université de Sherbrooke)

Mehmet Yayla (non affilié, précédemment à Pêche et Océans Canada et à l'Institut Maurice-Lamontagne)

Résumé/abstract

La prolifération des cyanobactéries est un problème croissant dans les milieux lacustres qui engendrent une forte dégradation de la qualité de l'eau. Certaines espèces produisent même des toxines nocives pour la santé des êtres vivants. La phycocyanine (PC) est un pigment photosynthétique typique des cyanobactéries et il est possible d'étudier ces dernières par télédétection. Au moins trois algorithmes de la littérature tentent d'estimer la concentration en PC ([PC]) à l'aide de la télédétection, mais leurs erreurs relatives sont importantes, surtout pour les faibles concentrations. Dans cette étude, nous présentons les résultats d'un nouvel algorithme qui emploie la variabilité de deuxième ordre (anomalies) des spectres de réflectance de l'eau pour estimer la [PC]. Cette méthode n'a jamais été employée dans les milieux lacustres. Le jeu de données utilisé pour développer et valider l'algorithme fut acquis entre 2001 et 2005 sur 57 lacs et réservoirs différents d'Espagne et des Pays-Bas. La performance de l'algorithme de deuxième ordre est équivalente sinon meilleure que celle des trois autres algorithmes déjà publiés. Pour le sous-ensemble de données où $[PC] > 32 \text{ mg m}^{-3}$, l'apport du terme de deuxième ordre ($R^2=0,68$ et $\text{RMSE}=0,25$) améliore significativement l'algorithme de premier ordre ($R^2=0,50$ et $\text{RMSE}=0,35$). L'exactitude des estimations de l'algorithme de 2^e ordre pour $[PC] > 32 \text{ mg m}^{-3}$ est supérieure à celle calculée pour l'ensemble du jeu de données ($R^2=0,69$ et $\text{RMSE}=0,44$). L'algorithme de deuxième ordre peut aussi s'adapter aux bandes du capteur satellitaire MERIS pour l'étude des cyanobactéries. L'application de cet algorithme à une image MERIS du 29 août 2010 au-dessus de la Baie Missisquoi (Québec, Canada) démontre le potentiel de ce nouvel algorithme pour un futur système de surveillance des cyanobactéries. Notons que tous les résultats statistiques présentés ci-haut sont calculés pour le logarithme de [PC] et que les unités du RMSE sont en $\log(\text{mg/m}^3)$.

Proliferation of cyanobacteria is a growing problem in lacustrine environment that results in rapid degradation of water quality. Moreover, certain cyanobacteria species produce harmful toxins. Phycocyanin (PC) is a photosynthetic pigment typical of cyanobacteria and affects the water color: it is therefore possible to study them using remote sensing. At least three algorithms to estimate PC concentration ([PC]) have been published, but their relative errors are important, especially for lower concentration. In this study, we are presenting the results of a new algorithm that uses the second order variability (anomalies) of water's reflectance spectrum to estimate [PC]. This method has never been used in lacustrine environment. The dataset used to develop and validate the algorithm was obtained between 2001 and 2005 in 57 different lakes and reservoirs of the Netherlands and Spain. The performance of the second order algorithm is equivalent or better than the three previously published algorithms. For the subset where $[PC] > 32 \text{ mg m}^{-3}$, the contribution of the second order term ($R^2=0.68$ and $\text{RMSE}=0.25$) seems to improve considerably the first order algorithm ($R^2=0.50$ and $\text{RMSE}=0.35$). The accuracy of the second order algorithm for $[PC] > 32 \text{ mg m}^{-3}$ is superior to the one calculated for the whole dataset ($R^2=0.69$ and $\text{RMSE}=0.44$). The algorithm can also be adapted to the bands of satellite sensor MERIS for the study of cyanobacteria. The application of this algorithm to a MERIS image acquired the 29 August 2010 taken over the Missisquoi Bay (Quebec, Canada) demonstrates the potential of this new algorithm for a future cyanobacteria' monitoring system. Note that all the statistical results presented above are for the logarithm of [PC] and the units of the RMSE are $\log(\text{mg/m}^3)$.

Mots-clés

Cyanobactéries, algues bleues, efflorescences toxiques, télédétection, phycocyanine, chlorophylle, lacs, variabilité de deuxième ordre, qualité de l'eau, MERIS.

Table des matières

Table des matières	4
Liste des annexes	5
Remerciements	6
1. Introduction	7
2. Problématique des cyanobactéries et de la télédétection	9
2.1 La phycocyanine	11
2.2 Le phytoplancton	13
2.3 La matière organique dissoute colorée	14
3. Télédétection des cyanobactéries	16
3.1 Télédétection de la chlorophylle <i>a</i> pour l'étude des cyanobactéries	16
3.2 Télédétection de la phycocyanine	22
3.3 Approche par variabilité de 2 ^e ordre	27
4. Objectifs et hypothèses	28
4.1 Objectifs principal et secondaires	28
4.2 Hypothèses principale et secondaire	28
5. Méthodologie	29
5.1 Site à l'étude	29
5.2 Étapes méthodologiques	32
6. Présentation, analyse et interprétation des résultats	39
6.1 Algorithme de variabilité de 2 ^e ordre	39
6.2 Comparaison avec d'autres algorithmes	45
7. Discussion des résultats	52
8. Conclusion	59
9. Références	63

Liste des annexes

Annexe 1 – Algorithme semi-analytique développé par Simis et al (2005)	69
Annexe 2 – Correction de la réflexion de surface de l'éclairement du ciel	70
Annexe 3 – Démonstration de la formule de correction du ciel.....	73
Annexe 4 - Analyses spectrales supplémentaires sur les anomalies	76
Annexe 5 – Deux algorithmes de 2e ordre alternatifs	80
Annexe 6 – Relation entre les erreurs relatives et le rapport des concentrations en pigments	84

Remerciements

Ce projet de recherche n'aurait jamais pu se concrétiser sans le soutien extraordinaire de quelques personnes et organismes. Lors des trois dernières années, je me suis toujours vanté d'avoir les meilleurs directeurs de recherche. Je veux tout d'abord remercier mon directeur Yannick Huot, pour avoir été constamment disponible et généreux de son temps. Malgré son jeune âge, il a su me guider avec sagesse tout au long de ma maîtrise et sans lui, je n'aurais pas terminé cette recherche avec autant de succès. Je souhaite aussi remercier mon codirecteur, Norm O'Neill, pour avoir cru en moi dès le début de cette aventure. Sa grande expérience, ses nombreuses connaissances et son sens de l'humour ont été très précieux pour moi. Guy Létourneau d'Environnement Canada m'a toujours soutenu à plusieurs niveaux : campagnes terrains au Québec, superviseur de stage à l'été 2009, financement, appuie dans ma revue de littérature, etc. Un grand merci à Guy et à toute l'équipe du Centre St-Laurent à Montréal (équipe terrain, techniciens de laboratoire, etc.). Je suis extrêmement reconnaissant aussi pour Stefan Simis du *Finnish Environment Institute* (SYKE) pour tout son aide et son temps en juin 2010 lorsqu'il m'a accueilli en Finlande. Il m'a beaucoup appris sur la télédétection des cyanobactéries et c'est grâce à lui et ces collègues de la Finlande et de l'Espagne que j'ai obtenu le jeu de données principal du présent ouvrage. Je veux remercier aussi l'Université de Sherbrooke et le Département de Géomatique appliquée pour offrir ce programme de recherche avec une belle équipe de professeurs (es), de professionnels (les) et d'employés de soutien. De plus, cette recherche fut en grande partie financée par le CRSNG, le FQRNT et le Cartel, et j'apprécie énormément leur soutien à la recherche. Finalement, je veux remercier ma famille et mon entourage pour m'avoir supporté toutes ses années!

1. Introduction

Le phénomène des blooms (efflorescences algales) toxiques de cyanobactéries, aussi appelé algues bleu-vert, atteint plusieurs lacs du Québec (150 plans d'eau en 2010 selon MDDEP, 2011) et du Canada chaque année (Wynne *et al.*, 2008). Cette problématique touche la santé des êtres vivants et toutes activités récréatives (Lavoie *et al.*, 2007). Dans la littérature, un manque d'études effectuées au Québec est observable dans le domaine de la télédétection des blooms de cyanobactéries. La télédétection désigne en fait l'acquisition à distance, sans contact direct, d'informations d'un objet, d'un territoire ou d'un phénomène. Étant donné l'étendue du territoire et le nombre de lacs à surveiller, l'instauration d'un programme de surveillance de ces derniers à l'aide de capteurs satellitaires ou aéroportés serait très pertinente, car il permettrait l'étude de la distribution spatiale des blooms ce qui est très difficile avec des mesures ponctuelles sur le terrain (Simis *et al.*, 2006; Wynne *et al.*, 2008).

La télédétection du phytoplancton (ensemble des organismes photosynthétiques du plancton en suspension dans l'eau) est possible par l'étude des caractéristiques du spectre de réflectance de l'eau. Puisque la réflectance de la lumière est inversement proportionnelle à l'absorption de l'eau et de ses composantes, il est possible d'estimer les concentrations de différents pigments en la mesurant. Les pigments présents dans le phytoplancton (Jeffrey et Vesk, 1997) provoquent l'absorption d'une partie de la lumière dans des bandes spectrales bien définies du visible (Mobley, 1994). Par exemple, la chlorophylle *a* (Chl *a*), à l'intérieur des cellules phytoplanctoniques, montre deux maximums d'absorption dans le bleu et dans le rouge, soit à 440 et 676 nm respectivement (Bricaud *et al.*, 2004). Puisqu'elle est présente en suspension dans l'eau uniquement dans le phytoplancton et que sa concentration est proportionnelle à la biomasse, la Chl *a* est généralement utilisée pour mesurer l'abondance en phytoplancton. Les cyanobactéries sont un des groupes formant le phytoplancton. On en compte environ 300 espèces au Québec et plusieurs ont la capacité de produire des toxines (Lavoie *et al.*, 2007). Elles synthétisent la Chl *a* ainsi que certains pigments accessoires comme la phycocyanine (*ibid.*). Dans les lacs eutrophes, des efflorescences algales, pouvant aller jusqu'à la formation de fleurs d'eau de cyanobactéries peuvent se développer très rapidement ce qui abaisse grandement la qualité de l'eau (Simis *et al.*, 2005; Lavoie *et al.*, 2007).

La phycocyanine (PC) est un pigment photosynthétique typique des cyanobactéries présente chez aucune autre espèce phytoplanctonique formant des fleurs d'eau. Il représente ainsi le meilleur indicateur optique de leur présence en raison de son pic d'absorption *in vivo* près de 620 nm (Simis *et al*, 2005). Lorsque les cyanobactéries sont en quantité suffisante, l'absorption par la PC crée un creux observable dans le spectre de réflectance entre 600 et 650 nm (Dekker, 1993; Schalles et Yacobi, 2000; Simis *et al*, 2005). L'absorption par la Chl *a* et de certains pigments accessoires ainsi que les effets d'absorption et de diffusion des autres composantes optiques de l'eau autour de 620 nm viennent se superposer à l'absorption causée par la PC ce qui complique l'estimation de la concentration des cyanobactéries à l'aide de la PC. Quelques algorithmes de télédétection ont déjà été développés dans le but de quantifier la PC (Dekker, 1993; Schalles et Yacobi, 2000; Simis *et al*, 2005). Malgré leur approche différente, il a été démontré que leurs erreurs relatives sont plus élevées pour les faibles concentrations ($PC < 50 \text{ mg/m}^3$) et les fortes concentrations ($PC > 200 \text{ mg/m}^3$) en PC (Ruiz-Verdú *et al*, 2008).

Brown *et al*. (2008) ont développé une nouvelle méthode en milieu océanique qui emploie la variabilité de 2^e ordre (les anomalies) des spectres de réflectance de l'eau afin d'estimer la contribution des paramètres optiques dont la contribution au signal est faible par rapport au constituant plus important (1^{er} ordre). Nous référons à ces constituants comme ceux de "2^e ordre". Cette démarche n'a jamais été appliquée en milieu lacustre, mais nous croyons qu'il existe un potentiel tangible pour le développement d'un algorithme d'estimation de la concentration en PC ([PC]) grâce à l'analyse des anomalies causées par la PC. Notre objectif principal sera donc de développer une méthodologie employant la variabilité de 2^e ordre de la PC pour l'étude des blooms de cyanobactéries.

2. Problématique des cyanobactéries et de la télédétection

L'eutrophisation des étendues d'eau, soit l'enrichissement naturel ou artificiel de l'eau en matières nutritives, est grandement accélérée par les activités humaines. Entre autres, le lessivage des terres agricoles, le ruissèlement accéléré par les routes et le traitement inadéquat des eaux usées précipitent la modification et la dégradation des masses d'eau (WHO, 2003). De plus en plus commune au Québec, cette problématique entraîne la prolifération des algues et des cyanobactéries dans les lacs (Lavoie *et al.*, 2007). Les cellules de cyanobactéries, contrairement à la plupart des algues, peuvent s'accumuler à très forte densité à la surface et former des fleurs d'eau (forme de l'efflorescence algale condensée à la surface) [WHO, 2003]. Les blooms de cyanobactéries surviennent principalement durant la période estivale (Lavoie *et al.*, 2007) ce qui correspond au moment où la demande en activité récréative sur les étendues d'eau est très forte. De plus, certaines espèces de cyanobactéries sont toxiques (Ernst *et al.*, 2005), amplifiant le besoin d'instaurer un système de surveillance de ces dernières (Environnement Canada, 2011). Plusieurs cas d'animaux morts par les cyanotoxines, comme des poissons, chiens, chats, cochons, vaches, moutons, zèbres et rhinocéros, ont été constatés (Lavoie *et al.*, 2007; WHO, 2003). Les humains ne sont pas plus à l'abri des algues bleu-vert et de leurs toxines (Lavoie *et al.*, 2007; WHO, 2003). Ces dernières sont à l'origine de plusieurs malaises (gastro-entérites, nausées, vomissement, fièvre, irritation des yeux et des oreilles, dommages aux reins et au foie, etc.) et décès (Lavoie *et al.*, 2007; WHO, 2003).

Puisque les cyanobactéries se retrouvent souvent près de la surface de l'eau grâce à leurs vacuoles gazeuses (Lavoie *et al.*, 2007), la télédétection pourrait être un outil complémentaire très intéressant par rapport aux méthodes traditionnelles d'échantillonnage. Par exemple, l'utilisation de la télédétection satellitaire aurait l'avantage de faciliter l'étude de la distribution spatiale des blooms sur de vastes territoires ce qui est très difficile avec des mesures ponctuelles sur le terrain (Simis, 2006; Wynne *et al.*, 2008). De plus, lors d'épisodes de blooms, la concentration des cyanobactéries varie rapidement dans le temps et même spatialement sur de courtes distances (possiblement en quelques heures et sur quelques mètres). La prise d'échantillons sur le terrain devient donc inefficace pour évaluer l'évolution de la population de phytoplanctons (Reinart et Kutser, 2006; Kutser *et al.*, 2006). En fait, les capteurs satellitaires permettent de surveiller les blooms de façon routinière et pratiquement en temps réel (Reinart et Kutser, 2006) selon un cycle

régulier de 0,5 à 6 jours dépendamment du capteur satellitaire et de la nébulosité (ICES, 2006). Avec la forte augmentation des épisodes de fleurs répertoriés sur les lacs du Québec depuis 2004 (Lavoie *et al.*, 2007), la télédétection des cyanobactéries apparaît non seulement comme un outil scientifique fort intéressant, mais désormais primordiale pour l'environnement et la santé et la sécurité des êtres vivants.

Tout de même, plusieurs défis technologiques, scientifiques et pratiques sont à surmonter. La détection à l'aide de capteurs satellitaires n'est possible que lorsqu'il y a absence de couverture nuageuse et que les connaissances et les outils pour éliminer les effets optiques de l'atmosphère (autre que les nuages) sont disponibles. Il faut que les cyanobactéries soient présentes près de la surface de l'eau, ce qui n'est pas toujours le cas (Lavoie *et al.*, 2007). Aussi, à l'aide de mesures spectroradiométriques uniquement (c'est-à-dire des mesures satellitaires de haute résolution spectrale), il est impossible de distinguer les différentes espèces entre-elles ainsi que les effets optiques des cyanobactéries des effets optiques de certaines propriétés de l'eau dont la turbidité (ICES, 2006). Puis, les blooms apparaissent souvent dans les eaux douces, les eaux côtières ou près des littoraux, à l'endroit même où la résolution spatiale plutôt grossière des satellites pose problème (ICES, 2006; Simis *et al.*, 2006). Dans bien des cas, la concentration en cyanobactéries à la surface est insuffisante pour être détecté de l'espace, mais assez importante pour être nuisible (ICES, 2006). De plus, il est primordial d'évaluer la toxicité d'une étendue d'eau et celle-ci se fait que par des mesures *in situ*. En fait, les différentes espèces de cyanobactéries ne sont pas toutes toxiques et leur toxicité dépend des conditions de croissance (Lavoie *et al.*, 2007).

La problématique de la télédétection des cyanobactéries a fait l'objet de plusieurs études scientifiques et malgré certaines contraintes, un fort potentiel existe pour son application. Par contre, avant d'analyser différentes études portant directement sur la détection des cyanobactéries, il est important de connaître certaines caractéristiques optiques des eaux naturelles qui modifient la réponse du signal.

La réflectance (R , sans unité) est définie comme étant le rapport entre l'éclairement ascendant provenant de l'eau (E_u , $W\ m^{-2}\ nm^{-1}$) et l'éclairement descendant provenant du ciel (E_d , $W\ m^{-2}\ nm^{-1}$). Les valeurs spectrales de réflectance ($R(\lambda)$) sont reliées aux constituants de l'eau, par

exemple les pigments contenus dans l'eau (Morel, 1988). Lorsque le rapport entre b_b/a est $\ll 1$, où b_b (m^{-1}) est le coefficient de rétrodiffusion total et " a " (m^{-1}) est le coefficient d'absorption total de l'eau, la réflectance de l'eau peut-être exprimée (ex. Morel, 1988) par :

$$R = 0,33b_b/a \quad (1)$$

Ainsi, l'absorption de la lumière par différents constituants comme l'eau pure, les pigments, la matière organique dissoute colorée et la matière en suspension vient influencer le signal provenant des étendues d'eau affectées par les algues bleu-vert et ils doivent être considérés avant le développement d'un algorithme de télédétection (voir Mobley [1994], pour une description des paramètres b_b et a). Dans les prochains paragraphes, la phycocyanine, le phytoplancton et la matière organique dissoute colorée (*yellow substance*) seront examinés de plus près. Les effets optiques de l'eau en-soi et la matière non algale en suspension ne sont pas négligeables non plus, mais elles ne seront pas présentées en détail ci-dessous. À noter que la mesure prise par un capteur de télédétection est la luminance (le rayonnement directionnel, voir Mobley 1994). Donc, on ne peut pas calculer la réflectance directement telle que définie ci-dessus, mais plutôt par un rapport entre la luminance sur l'éclairement descendant. La réflectance de télédétection est en fait un rapport d'une luminance sur un éclairement descendant. Elle sera définie plus en détail dans la section 2.5 Méthodologie.

2.1 La phycocyanine

De façon générale, les algues possèdent des pigments absorbant l'énergie solaire tout comme la cyanobactérie qui est une bactérie photosynthétique faisant partie du phytoplancton. Le processus de la photosynthèse nécessite des pigments comme la Chl *a* contenue dans tous les végétaux (Mobley, 1994). Par contre, les cyanobactéries utilisent en plus la PC, une phycobiliprotéine, afin de réaliser la photosynthèse (Lavoie *et al*, 2007). Ce pigment bleuâtre présente un pic d'absorption à environ 620 nm et il varie quelque peu selon les espèces (Simis, 2005). La Figure 1 de Farabee (2009) montre le spectre d'absorption de quelques pigments dont la PC (*phycocyanin* en anglais). Il est possible d'y observer que la largeur à mi-hauteur du pic (ou FWHM pour *full width at half maximum* qui représente la largeur du pic d'absorption au niveau de la moitié de la valeur du pic) est assez large pour le pic d'absorption de la PC à 620 nm (de ~

60 à 70 nm vs ~ 20 nm pour le Chl *a*). Par contre, il est important de noter qu'une certaine incertitude est présente dans la littérature par rapport au FWHM de la PC. C'est l'emplacement du maximum d'absorption à 620 nm, par rapport aux autres groupes du phytoplancton, qui permet de détecter distinctement les algues bleu-vert par télédétection. De plus, notons que la longueur d'onde d'émission de la PC est 650 nm lors de la fluorescence (Lavoie *et al*, 2007). La fluorescence est en fait un processus d'émission lumineuse spontané à une longueur d'onde donnée provoquée par l'absorption de la lumière à une autre longueur d'onde.

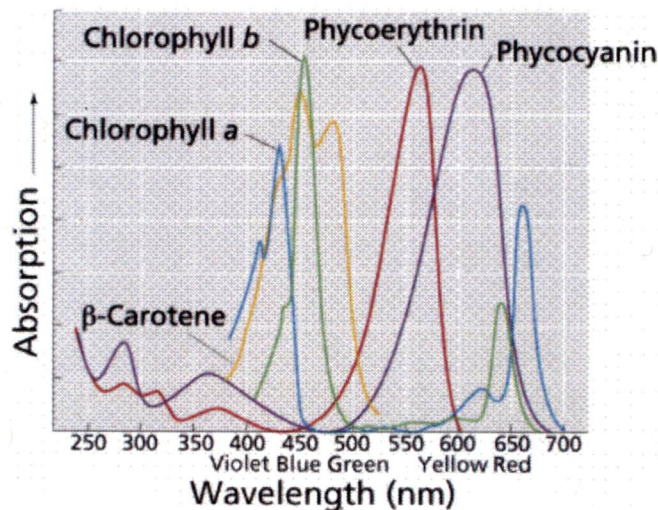


Figure 1 : L'auteur ne donne pas de détails, mais l'étiquette 'Absorption' doit référer au coefficient d'absorption des différents pigments (dont la phycocyanine) en fonction de la longueur d'onde (Farabee, 2009).

Dans une étude menée par Kutser *et al* (2006), il est mentionné que les cyanobactéries analysées possèdent un coefficient d'absorption maximal à 631 nm, mais que cette absorption n'est pas uniquement due à la concentration en PC. *In vitro*, ce pic d'absorption serait entre 615 et 621 nm selon les mêmes auteurs. Toujours dans cette même étude, les auteurs ont observé que le pic d'absorption de la PC dans certains lacs des Pays-Bas est à 627 nm. De ce fait, nous croyons que d'une étendue d'eau à l'autre, le pic du maximum d'absorption observé de façon *in vivo* peut varier quelque peu (généralement entre 615nm et 630nm d'après la littérature), donc il serait normal de voir une variance à ce niveau sur les différents lacs du Québec ou du Canada. En plus, selon Simis *et al* (2005), le coefficient d'absorption spécifique de la PC (qui normalise les variations beaucoup plus importantes dues à la variation de la concentration) est variable, donc il varie selon la taille et la nature de l'algue bleu-vert.

Notons aussi qu'il n'y a aucune méthode standard à ce jour pour l'extraction de la concentration de la PC (c'est-à-dire les mesures *in vivo* ou *in situ* qui serviraient de vérité de terrain pour les mesures de télédétection). Simis *et al.* (2006) ont utilisé dans leur étude deux méthodes différentes pour l'extraction de la concentration de la PC en laboratoire (le cycle répété de gel et dégel et le broyage mécanique) et une méthode de mesure *in situ* de la PC par fluorescence afin de comparer leur performance. Les trois méthodes explorées donnent des résultats assez similaires (c'est-à-dire que la pente d'une méthode par rapport à l'autre est proche de 1), mais avec beaucoup de dispersion autour de la courbe de régression. Donc, pour l'instant, il faut considérer les mesures de vérité terrain avec précautions. La prise de mesures *in situ* de PC par fluorescence est maintenant populaire et prometteuse selon Simis *et al.* (2006), car elles sont faites facilement, sans étape d'extraction et d'échantillon à préserver.

2.2 Le phytoplancton

Les propriétés d'absorption des pigments sont importantes dans les eaux naturelles. En effet, plusieurs pigments présents dans le phytoplancton (Jeffrey, 1997) provoquent l'absorption d'une partie du spectre lumineux dans le visible (Mobley, 1994). La Chl *a* est l'un des principaux pigments, car il est présent dans tous les végétaux et il est employé pour mesurer l'abondance en phytoplancton. La Chl *a* possède deux maximums d'absorption, soit dans le bleu et dans le rouge, à 440 et 676 nm respectivement (Bricaud *et al.*, 2004). La concentration de la Chl *a* ([Chl *a*]) varie d'environ 0,01 mg/m³ dans les eaux océaniques très claires à 100 mg/m³ dans les lacs eutrophes (Mobley, 1994). Reinart et Kutser (2006) prétendent même que les concentrations vont au dessus de 100 mg/m³ durant les épisodes de blooms dans la mer Baltique. Nous constatons le même phénomène dans certains lacs d'Espagne étudiés par Simis *et al.* (2006) où des concentrations au-dessus de 400 mg/m³ ont été enregistrées. Les études empiriques (Bricaud *et al.*, 1983) et théoriques (Morel et Bricaud, 1981) ont montré que le spectre d'absorption du phytoplancton peut varier grandement selon ses caractéristiques cellulaires et en particulier sa taille (Bricaud *et al.*, 2004). Ces variations sont causées par le *pigment packaging effect* ("discrétisation" des pigments dans la cellule, qui varie elle-même selon l'espèce, les conditions environnementales et de la grosseur des cellules) et la composition des pigments (Chl *b*, Chl *c*, etc.). Pour illustrer cette variabilité, la Figure 2 suivante illustre 8 spectres de coefficient

d'absorption spécifique de la Chl provenant de 8 espèces de phytoplanctons différentes ($a^*_{\text{Chl}} = a_{\text{Chl}} / [\text{Chl}]$).

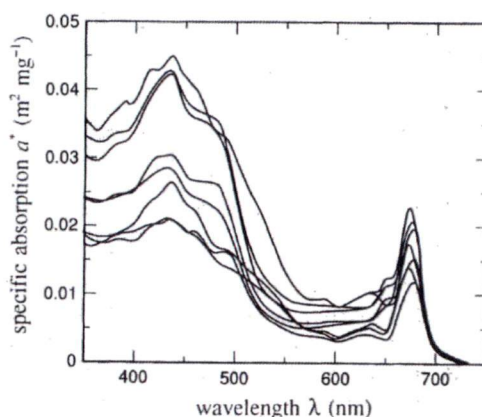


Figure 2 : Spectres de coefficient d'absorption spécifique pour 8 espèces de phytoplanctons normalisée à la concentration de chlorophylle (Mobley, 1994).

Le coefficient d'absorption spécifique est par définition indépendant de la quantité de chlorophylle dans la culture et il varie donc seulement selon la taille et la nature des cellules étudiées. La Figure 2 montre que la position du pic d'absorption dans le rouge (centrée près de 676 nm) peut varier d'environ 10 nm. Ce phénomène est généralement causé par la contribution d'autres pigments (en particulier la Chl *b*; Bricaud *et al*, (1983)). Enfin, le FWHM est environ de 15 nm (*ibid.*).

Un autre phénomène optique important à considérer pour la Chl *a* est la fluorescence. La longueur d'onde d'excitation des pigments chlorophylliens qui génère une émission maximum de fluorescence autour de 685 nm est presque identique à celle de leur maximum d'absorption, donc à 440 nm pour la Chl *a* (Mobley, 1994; American Public Health Association, 2005). L'indépendance de la longueur d'onde d'émission en fonction de la longueur d'onde d'excitation est une caractéristique des substances pures, soit une substance chimique caractérisée par une composition et des propriétés spécifiques (Mobley, 1994). Ce phénomène non négligeable vient augmenter la réflectance observée dans le rouge autour de 685 nm.

2.3 La matière organique dissoute colorée

La matière organique dissoute colorée (CDOM), aussi connue sous les noms de *yellow substance*, *CDOM* ou *Gelbstoff*, est la composante mesurable optiquement provenant de la décomposition de

la matière organique (NASA Science, 2009). En milieu côtier, il provient principalement des acides humiques et fulviques et de la dégradation du phytoplancton (Froidefond et Doxaran, 2004). Le CDOM absorbe fortement la lumière dans les courtes longueurs d'onde (UV, bleu). Puis, son coefficient d'absorption diminue de façon exponentielle jusqu'à atteindre une valeur négligeable dans le PIR (voir Figure 3). Les eaux non turbides, avec très peu de CDOM, ont une couleur bleue à cause de l'absorption par les molécules d'eau dans le rouge combiné avec la rétrodiffusion des constituants (la couleur exacte étant une fonction du niveau de turbidité). Par contre, une eau possédant une forte quantité de CDOM a une apparence plutôt jaune-verte ou même brune (Froidefond et Doxaran, 2004). Selon Mobley (1994), la forme exponentielle suivante décrit le coefficient d'absorption du COD pour λ entre 350 et 700 nm :

$$a_y(\lambda) = a_y(\lambda_0) \exp[-0,014(\lambda - \lambda_0)] \quad (2)$$

La longueur d'onde de référence, λ_0 (nm), est à 440 nm et $a_y(\lambda_0)$ représente le coefficient d'absorption du CDOM à 440 nm. Toujours selon Mobley (1994), $a_y(\lambda)$ dépend bien évidemment de sa concentration dans l'eau et de sa nature. L'auteur mentionne aussi que l'expression ci-dessus permet d'estimer le spectre d'absorption du CDOM si son coefficient d'absorption est connu à une longueur d'onde, mais qu'il n'y a toujours pas de modèle qui permet d'obtenir le spectre d'absorption du CDOM si seulement sa concentration est connue (autrement dit, le coefficient d'absorption spécifique du CDOM, $a_y(\lambda) / [y]$ n'est pas bien connu).

Selon Morel et Gentili (2009), $a_y(440)$ est faible en milieu océanique ($0,010$ à $0,113 \text{ m}^{-1}$) alors que sur les côtes et littoraux, elle peut varier entre $0,07$ et $0,65 \text{ m}^{-1}$ jusqu'à atteindre des valeurs près de 20 m^{-1} . La Figure 3, tirée de Froidefond et Doxaran (2004), illustre un exemple de $a_y(\lambda)$, pour $a_y(\lambda_0) = 0,6 \text{ m}^{-1}$.

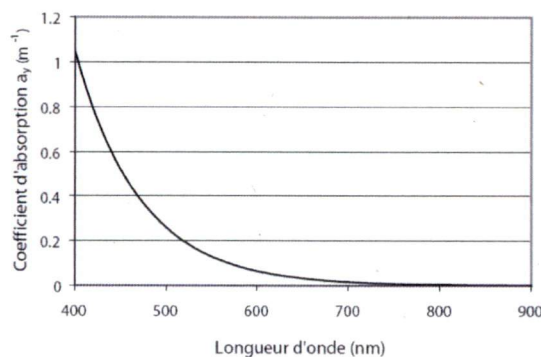


Figure 3 : Variation du coefficient d'absorption du CDOM en fonction de la longueur d'onde définie pour $a_y(\lambda_0) = 0,6 \text{ m}^{-1}$.

3. Télédétection des cyanobactéries

Maintenant que l'on comprend mieux l'effet optique de certains constituants de l'eau, nous pouvons nous concentrer sur la télédétection des cyanobactéries. En forte concentration, c'est la PC qui donne aux cyanobactéries la couleur bleuâtre. De plus, grâce à leur vacuole gazeuse, les algues bleu-vert ont la capacité de se déplacer verticalement dans l'eau en période calme pour absorber un maximum de lumière près de la surface durant le jour (Lavoie *et al*, 2007). Les blooms ou les fleurs d'eau de cyanobactéries se forment lorsque ces dernières croissent abondamment entraînant une augmentation importante de la biomasse (*ibid.*). De plus, il est parfois possible d'observer une écume flottante de cyanobactéries lorsqu'elles prolifèrent ou s'accumulent à la surface de l'eau (*ibid.*). Ainsi, l'étude des cyanobactéries par télédétection est réalisable grâce à leurs particularités spectrales distinctes (absorption de la PC et Chl *a* par exemple) et à leur présence près ou à la surface de l'eau (donc possibilité de les observer de l'espace, car la pénétration effective de l'éclairement solaire, et donc la profondeur à laquelle les capteurs satellitaires sont sensibles, est seulement de quelques mètres pour les eaux turbides).

De nombreuses études se sont penchées sur la télédétection des cyanobactéries. Certaines s'attaquent aux algorithmes de détection de la cyanobactérie (Vincent *et al*, 2004, Kutser *et al*, 2006), de son pigment la PC (Dekker, 1993; Schalles et Yacobi, 2000; Simis *et al*, 2005; Simis *et al*, 2006) ou bien aux concentrations importantes de Chl *a* (Réinart et Kutser, 2006; Doerffer et Schiller, 2007; Gower et King, 2007; McKee *et al*, 2007). D'autres projets de recherche ont pour objectif d'étudier la capacité et l'efficacité des capteurs spectraux à surveiller les blooms de façon régulière pour l'instauration de système de surveillance fonctionnel (ICES, 2006; Park et Ruddick, 2007; Wheeler *et al.*, in press). Trois grands thèmes seront abordés dans cette section : la télédétection de la Chl *a* pour l'étude des cyanobactéries, la télédétection de la PC et l'approche par variabilité de 2^e ordre.

3.1 Télédétection de la chlorophylle *a* pour l'étude des cyanobactéries

La Chl *a* est bien souvent employée comme variable dans plusieurs études liées à la détection de blooms de cyanobactéries, partiellement en raison du faible échantillonnage de la PC comparativement à la Chl *a* mesurée de façon régulière dans les programmes de surveillance (Kutser *et al*, 2006). Même si la Chl *a* que l'on détecte associée à la PC n'est pas un indicateur

direct de la présence de PC (la majorité des phytoplanctons ont la Chl *a*, mais aucune PC), nous croyons qu'il est intéressant d'étudier sa concentration comme indicateur potentiel de la présence des cyanobactéries. En fait, cette donnée nous informe sur la biomasse algale contenue dans l'étendue d'eau analysée et elle est un bon complément au nombre limité d'études s'intéressant uniquement aux propriétés propres aux algues bleu-vert.

Plusieurs algorithmes empiriques emploient des bandes ou des rapports de bandes de réflectance afin de déterminer la [Chl *a*]. Matthews *et al* (2010) ont justement analysé et comparé plusieurs algorithmes empiriques lors d'une étude menée en Afrique du Sud dans un lac hypereutrophe peu profond (lac Zeekoevlei). Les auteurs ont, entre autres, étudié la performance d'algorithmes empiriques pour l'estimation de la [Chl *a*] et de la quantité totale de matière en suspension (MES). Les mesures spectroradiométriques ont été acquises de façon *in situ* et à partir du capteur MERIS. Il est important de savoir que la méthode empirique ne tient pas compte de l'optique fondamentale de l'eau et les relations statistiques peuvent être faussées par l'influence de constituants corrélés, alors elle doit être utilisée avec précaution (Mouchot *et al*, 1996). De plus, l'application d'un modèle empirique est souvent limitée à la zone d'étude employée pour développer l'algorithme. Dans l'étude de Matthews *et al* (2010), certains algorithmes empiriques donnent des résultats similaires pour différents paramètres ([Chl *a*], MES, etc.). La forte covariance entre ces paramètres et le fait qu'il n'y a pas de bandes uniques pour l'étude du phytoplancton et la MES par exemple explique cette constatation (Matthews *et al.*, 2010). En raison de la nature empirique des algorithmes étudiés dans Matthews *et al* (2010), nous ne pouvons pas conclure que leur performance sera similaire dans d'autres milieux comme au Québec, à moins que les propriétés des données soient sensiblement les mêmes comme la forte covariance entre certains paramètres.

Néanmoins, les résultats de certains algorithmes empiriques sont forts encourageants. Par exemple, le meilleur algorithme ($R^2=0,925$) pour estimer la [Chl *a*] selon les analyses de Matthews *et al* (2010) est le simple rapport de bande entre R708/R664 lorsqu'il est appliqué à des données de MERIS acquises au-dessus de l'atmosphère. Ces résultats confirment le fort potentiel de MERIS pour l'emploi d'algorithmes très simples sur des lacs eutrophes sans l'utilisation de corrections atmosphériques qui posent souvent problème en milieu lacustre. En général, les

données de réflectance acquises aux dessus de l'atmosphère (sans correction atmosphérique) donnent de meilleurs résultats lorsqu'elles sont appliquées à des algorithmes de rapports de bandes (Matthews *et al.*, 2010). Ce phénomène s'explique par le fait que les rapports de bandes normalisent approximativement les effets de l'atmosphère, pourvu que les bandes soient relativement proche spectralement (R709/R664 par exemple). Par contre, les algorithmes à bande unique performant mieux pour des données qui ont été corrigées pour les effets de l'atmosphère (ibid.).

Gitelson *et al.* (1993) ont aussi étudié les bandes qui sont les plus sensibles aux variations de [Chl *a*] à l'aide d'une analyse en composantes principales. Les auteurs ont découvert que la bande à 700 nm était la plus sensible aux variations de [Chl *a*] pour leur jeu de données. Ils ont développé deux algorithmes, un employant le rapport R700/R560 et l'autre R700/R675, afin d'estimer avec précision la [Chl *a*] ($R^2 = 0,92$ et $0,989$ respectivement sur le lac Müggelsee en Allemagne). Les deux algorithmes ont offert des résultats forts intéressants, particulièrement celui qui emploie le rapport R700/R675. Ce résultat est très similaire à celui rapporté par Matthews *et al.* (2010) et les deux bandes du deuxième algorithme (675 et 700 nm) sont relativement proche des bandes 7 et 9 de MERIS respectivement centrées à 665 et 705 nm. Dans le même ordre d'idée, Dekker (1993), Gons (1999) et Gons *et al.* (2000) ont tous démontré que le rapport de bande entre R708/R676 est un bon indicateur de [Chl *a*] dans les eaux intérieures eutrophes.

En fait, traditionnellement les algorithmes de télédétection de la Chl *a* ont été développés pour des eaux du cas I (milieu océanique) en utilisant les bandes dans la région du bleu et du vert (Morel et Prieur, 1977; Binding *et al.*, 2011). Par contre, en milieu lacustre où l'eau est plus turbide, l'apport de paramètres additionnels comme le CDOM dans ses longueurs d'onde vient fausser les prédictions de [Chl *a*]. De plus, les problèmes liés aux corrections atmosphériques sur les eaux turbides des eaux du cas II affectent davantage les bandes de courtes longueurs d'onde du visible (Moore *et al.*, 1999).

Pour répondre à ces nouvelles exigences, le capteur MERIS employé sur le satellite européen Envisat est bien adapté pour l'étude des blooms du phytoplancton grâce entre autres à sa bande à 709 nm qui n'est pas présente chez d'autres capteurs de la couleur de l'océan comme MODIS ou

SeaWiFS (Gower et King, 2008). En plus des produits pour l'étude des océans, le capteur MERIS génère une série de produits pour les eaux de cas II (eaux intérieures et côtières où les propriétés optiques ne covarient pas avec la [Chl *a*]). Par exemple, les produits *algal_2* pour la [Chl *a*], *yellow_subs* (CDOM) et *total_susp* (sédiments en suspendus) sont retrouvées à partir d'un modèle numérique de réseau de neurones appliqué aux données L2 de MERIS (Doerffer et Schiller, 2007). Plusieurs autres produits et options reliés à la biomasse phytoplanctonique sont offerts par le logiciel de traitement BEAM disponible gratuitement pour les utilisateurs de MERIS. Par exemple, les options suivantes sont disponibles : des produits qui diffèrent par leur correction atmosphérique adaptée aux conditions du milieu comme *Case_2 Regional* (C2R), un module, nommé *ICOL*, qui tente d'éliminer l'effet d'adjacence (la contribution aux effets atmosphériques de la lumière qui est réfléchi par des cibles et diffusée hors du champ de visé de chaque pixel) pour les lacs de petite taille, etc. (Binding *et al.*, 2011). Il y a aussi deux produits reliés à la Chl *a* qui sont offerts, soit la *Fluorescence line height* (FLH) ou le *Maximum Chlorophyll Index* (MCI). Une étude menée par Binding *et al.* (2011) sur Lake of the Woods avait pour objectif de tester l'applicabilité de ces produits de [Chl *a*] de MERIS pour l'étude des blooms d'algue-bleu sur une étendue d'eau optiquement complexe. La comparaison des produits *algal_2* et C2R avec les données *in situ* recueillies donne de mauvais résultats ($R^2=0,298$ et $R^2=0,255$ respectivement). Ces produits semblent incapables de prédire la présence ou l'intensité des blooms, même lorsque différents modules pour les corrections atmosphériques sont appliqués (Binding *et al.*, 2011). Le MCI est le seul produit relié à la [Chl *a*] à produire des résultats respectables ($R^2=0,739$). Le MCI calcul la luminance sortant de l'eau à 709 nm relativement à la ligne de base créée entre les luminances des bandes à 681 et 754 nm (Gower et King, 2008). Il permet de localiser la présence de concentration importante de Chl *a* à la surface (ibid.). Le pic près de 700 nm est majoritairement causé par l'interaction entre l'absorption de l'eau et la rétrodiffusion des cellules de phytoplancton à la surface de l'eau. Sur Lake of the Woods, les utilisateurs des produits de MCI de MERIS ont réussi à bien estimer l'intensité des blooms et sa variabilité spatiale puisqu'il semble bien performer malgré les difficultés qui perdurent au niveau des corrections atmosphériques (Binding *et al.*, 2011). Un potentiel existe donc pour l'implantation d'un système de surveillance à l'aide de cet algorithme, mais ceci reste à vérifier sur d'autres milieux possédants des propriétés optiques différentes. Par contre, un tel algorithme ne pourrait pas faire la distinction entre un bloom de cyanobactéries et un bloom d'algue ou

d'une autre espèce de phytoplancton, car il n'exploite pas la bande à 620 nm de MERIS qui correspond au maximum d'absorption de la PC. En ce qui a trait aux résultats du FLH, ces derniers présentaient une relation inverse avec les données *in situ* de [Chl *a*] (Binding *et al.*, 2011). Cet algorithme sera étudié plus en détail un peu plus loin dans cette section.

En plus du capteur MERIS, les capteurs satellitaires MODIS et SeaWiFS, étudiant la couleur de l'océan (*ocean color sensors*), sont aussi capables d'étudier l'abondance en phytoplancton par l'emploi d'images de concentrations en chlorophylle. Ces produits seraient une source d'information régulière et pertinente sur la biomasse phytoplanctonique si leurs estimations ont une bonne précision pour de faibles allant à de fortes concentrations en Chl *a* (dans le cas d'un bloom de cyanobactéries). Reinart et Kutser (2006) ont mené une étude qui compare les produits de chlorophylle provenant de MERIS, MODIS et SeaWiFS acquis pour une partie de la mer Baltique. Les valeurs obtenues par leur algorithme respectif furent comparées entre eux et avec les valeurs de [Chl *a*] obtenues à partir d'un modèle biooptique qui relie les spectres de réflectance du capteur hyperspectrale Hyperion à des spectres de référence variant selon la [Chl *a*].

Les algorithmes calculant la [Chl *a*] avec les trois capteurs étudiés donnent sensiblement les mêmes résultats au début des blooms, lorsque les algues sont bien distribuées dans la colonne d'eau (Reinart et Kutser, 2006). Par contre, de grandes différences sont observables pour des blooms très importants et situés à la surface de l'eau. En plus, il a été observé que tous les algorithmes ont tendances à sous-estimer la quantité réelle de Chl *a* durant les blooms de cyanobactéries (*ibid.*). Ceci pourrait être expliqué par la variabilité de la [Chl *a*] même à l'intérieur d'un pixel de fleur d'eau, l'effet d'entassement des pigments (*pigment packaging effect*) et aussi par les problèmes engendrés lors de la correction atmosphérique (*ibid.*). Il y a aussi la possibilité que les concentrations en Chl *a* des blooms de la mer Baltique furent au-dessus des limites de détection de MODIS, de SeaWiFS et de MERIS. Ces limites, n'étant pas bien répertoriées dans l'étude, devront peut-être faire l'objet de futures recherches pour connaître le véritable potentiel des produits de Chl de ses capteurs.

En raison de la configuration de leurs bandes, les capteurs SeaWiFS et MODIS n'ont pas la capacité d'étudier directement la [PC] et par le fait même la présence de cyanobactéries. Par

contre, si la prédominance des cyanobactéries dans la colonne d'eau est prouvée, nous croyons que plus il y aura de la Chl *a* détectée (indice de biomasse), plus il y aura de cyanobactéries. Ainsi, il sera alors possible d'associer l'importance de la concentration en Chl *a* à celle des cyanobactéries. Donc, en ce qui a trait à la télédétection des cyanobactéries, les capteurs MODIS et SeaWiFS sont utiles uniquement lorsque des analyses en laboratoire, des mesures terrains ou un capteur à meilleure résolution spectrale comme MERIS (bande à 620 nm) indiquent clairement que la biomasse estimée est bel et bien reliée à la présence de cyanobactéries. Leur principal attrait se situe au niveau de leur résolution temporelle. Par exemple, le capteur MODIS est présent à la fois sur le satellite Terra et Aqua ce qui permet la couverture d'un même endroit jusqu'à quatre fois par jour. Notons aussi que le capteur SeaWiFS a une résolution spatiale désavantageuse de 1000 m par rapport à MERIS (300 m de résolution) et MODIS (250 m pour les bandes B1 à 648 nm et B2 859 nm, 500 ou 1000 m pour les autres bandes) qui sont mieux adaptés pour la surveillance de lacs de moyenne taille (Reinart et Kutser, 2006). Malgré leur résolution spatiale relativement grossière et la configuration de leurs bandes, les satellites étudiant la couleur de l'océan sont présentement les seuls instruments disponibles pour surveiller les blooms de cyanobactéries de façon routinière et pratiquement en temps réel (ibid.).

La Chl *a* peut aussi s'étudier par la relation entre sa fluorescence à 685 nm et sa concentration. Un algorithme de ligne de base développé par Neville et Gower (1977) calcule la luminance en excès à la bande 8 de MERIS (681 nm) par rapport à la ligne de base formée entre les luminances de la bande 7 et 9 (665 et 709 nm). Une relation empirique est établie entre le signal de fluorescence retrouvé par la ligne de base (*Fluorescence Line Height*, FLH) et des mesures de [Chl *a*]. L'une des limites mise en évidence par les auteurs est que cette méthode n'est pas utilisable pour de fortes valeurs de [Chl *a*]. Pour les concentrations supérieures à 20 mg/m³, la méthode de la ligne de base sature. Une deuxième méthode devrait vraisemblablement être implantée lorsque la [Chl *a*] > 20 mg/m³ et cela augmente le travail et la complexité pour l'implantation d'un futur système de surveillance. Nous nous questionnons aussi si une seule formule est adéquate pour l'étude de tous les lacs d'un grand territoire, car la fluorescence dépend de la [Chl *a*], mais aussi du rendement quantique des différentes algues ou cyanobactéries. De plus, nous savons que la luminance varie avec l'intensité du soleil (soit de façon passive en fonction de la réflectance à 665 et 709 nm et, dans le cas spécial de fluorescence

à 681 nm, de façon active ainsi que passive). Alors, il y aura des valeurs de luminance plus élevées à 681 nm selon les saisons, l'heure et la nébulosité. Selon le même principe, la quantité de lumière absorbée par la Chl *a* variera aussi au cours de l'année et par le fait même la fluorescence variera de même en fonction de l'éclairement du soleil, et non pas seulement selon la [Chl *a*]. En raison de cette problématique, des valeurs de réflectance seraient probablement plus adéquates que les luminances pour l'étude de la [Chl *a*] par la fluorescence. Néanmoins, selon nous, cette méthode semble comporter plusieurs obstacles pour l'étude de blooms de façon routinière. Le Tableau 1-ci-dessous présente la liste des algorithmes de [Chl *a*] étudiés dans cette recherche.

Référence	Principe de l'algorithme
Matthews <i>et al.</i> (2010)	Rapport de bandes, soit R708/R664 (compatible avec MERIS).
Gitelson <i>et al.</i> (1993)	Rapport de bandes, soit R700/R560 ou R700/R675.
Dekker (1993); Gons (1999); Gons <i>et al.</i> (2000)	Rapport de bandes, soit R708/R676.
Gower et King (2007)	Algorithme FLH de ligne de base qui calcule la luminance en excès à 681 nm par rapport à la ligne de base entre 665 et 709 nm (compatible avec MERIS).
Gower et King (2008)	Le MCI calcul la luminance en excès à 709 nm relativement à la ligne de base entre les luminances à 681 et 754 nm (compatible avec MERIS).

Tableau 1 : Liste des algorithmes de [Chl *a*] présentés dans cette section.

La [Chl *a*] ne fait pas de lien explicite avec la concentration en cyanobactéries. Par contre, nous l'avons étudié afin de mieux comprendre la relation entre la Chl *a* et la PC et afin d'étudier les algorithmes principaux de détection de la Chl *a* qui dans un deuxième temps pourraient agir comme approximation de la [PC]. Maintenant, nous allons nous pencher sur la détection plus directe des cyanobactéries par l'entremise de son pigment typique, soit la PC.

3.2 Télédétection de la phycocyanine

Contrairement à la Chl *a*, l'analyse des caractéristiques spectrales de la PC permet d'étudier directement les cyanobactéries par télédétection. Effectivement, tel que mentionné plus tôt, la PC est un pigment unique aux cyanobactéries parmi les groupes de phytoplanctons formant des fleurs d'eau. Ainsi, l'absorption additionnelle de la lumière autour de 620 nm engendrée par la PC devrait s'observer uniquement en présence de cyanobactéries dans les masses d'eau à l'étude. L'un des obstacles toujours majeurs à l'étude de la PC est la grande difficulté à extraire ce dernier

en laboratoire. Aucune méthode standard n'a été proposée jusqu'à ce jour. Néanmoins, l'étude de la PC par télédétection a déjà été effectuée de différentes façons.

Étant donné que le maximum d'absorption de la PC est à 620 nm, quelques algorithmes empiriques et semi-empiriques examinant la réflectance dans le rouge ont été développés dans le passé. Les algorithmes de Dekker (1993) et Schalles et Yacobi (2000) furent les deux premiers publiés. Dekker (1993) tente d'étudier le creux de réflectance à 624 nm qui s'accroît à mesure que la [PC] augmente. Pour ce faire, il emploie deux autres bandes autour de 624 nm afin de former une ligne de base. Cette dernière est constituée des bandes à 600 et 648 nm qui sont près du maximum d'absorption de la PC sans toutefois en être affectées (Dekker, 1993). La réflectance moyenne entre les bandes à 600 et 648 nm soustraite à la réflectance à 624 nm est calculée pour estimer la [PC]. Notons que l'algorithme de Dekker (1993) a été conçu pour des réflectances acquises sous la surface de l'eau.

Schalles et Yacobi (2000) ont examiné des spectres de réflectance de l'eau en lien avec la concentration en pigments auxiliaires comme la PC et les caroténoïdes. Ils ont observé, tel que Dekker (1993), qu'en présence de PC un creux apparaît à 625 nm alors qu'un maximum de réflectance se crée à 650 nm en raison du minimum d'absorption local de la combinaison de PC et de la Chl *a* (voir la figure 1). Selon les auteurs, le simple rapport de bandes entre le maximum de réflectance à 650 nm et le minimum à 625 nm (R_{650}/R_{620}) causé par l'absorption de la PC leur donne la meilleure corrélation par rapport à leurs mesures de [PC] analysées en laboratoire ($R^2=0,612$). Les auteurs ont noté que la variabilité de la relation entre R_{650}/R_{620} et la concentration en PC était beaucoup plus forte pour les concentrations au-dessus de 200 mg/m³. Ils attribuent ce phénomène au problème qu'ils ont rencontré au niveau de l'extraction de la PC.

Les modèles semi-analytiques (les modèles qui combinent des expressions empiriques avec un modèle de transfert radiatif) peuvent aussi être utilisés pour l'étude de la biomasse de cyanobactéries puisque les algues bleu-vert ont des propriétés optiques inhérentes (IOPs) bien particulières et connues (voir Mobley, 1994, pour une définition des propriétés optiques inhérentes et apparentes). Simis *et al.* (2005) et (2006) ont tenté d'estimer la [PC] à partir, entre autres, des caractéristiques d'absorption de la PC à 620 nm. Plus particulièrement, leur

l'algorithme permet de retrouver la [PC] dans des eaux eutrophes et turbides à partir de rapport de bandes et de données sur les coefficients d'absorption spécifiques de la Chl *a* et de la PC dans le rouge et proche infrarouge. En vue d'utiliser leur algorithme sur les données MERIS, les auteurs utilisent les deux rapports de réflectance suivant : R709/R665 nm pour la Chl *a* et R709/R620 pour la PC (voir le Tableau 2 pour connaître les bandes de MERIS). Ces rapports servent à retrouver, à partir d'une simple application de l'expression de Gordon *et al* (1975), le coefficient d'absorption de la Chl *a* à 665 nm et celui de la PC à 620 nm en eau eutrophe. Les données qui ont servi au développement de l'algorithme proviennent de deux lacs hautement eutrophes, soit les lacs Loosdrecht et IJsselmeer des Pays-Bas. Il faut noter que l'algorithme contient certaines simplifications telles que mentionnées dans l'étude. Par exemple, l'absorption à 709 nm pouvant provenir en partie des pigments est considérée comme étant négligeable. De plus, l'absorption par la CDOM et le tripton (matière organique particulaire) n'est pas directement considérée non plus pour simplifier le calcul du coefficient d'absorption de la Chl *a* à 665 nm en presumant que son effet est suffisamment négligeable (Simis *et al.*, 2005). Par contre, un facteur de correction est employé pour éliminer l'effet de l'absorption de la Chl *a* à 620 nm, contrairement aux algorithmes de Dekker (1993) et de Schalles et Yacobi (2000). Pour plus de détails sur l'algorithme de Simis *et al.* (2005), veuillez vous référer à l'annexe 1.

No.	Centre de la bande (nm)	Largeur de la bande (nm)
1	412,5	10
2	442,5	10
3	490	10
4	510	10
5	560	10
6	620	10
7	665	10
8	681,25	7.5
9	708,75	10
10	753,75	7.5
11	760,625	3.75
12	778.75	15
13	865	20
14	885	10
15	900	10

Tableau 2 : Caractéristiques des bandes de MERIS. (Source: MERIS user guide issue 2.1, Table 1.1, <http://envisat.esa.int/envisat/dataproducts/meris/>).

Malgré certaines simplifications au niveau des paramètres pris en considération dans le modèle semi-analytique, l'algorithme présenté performe relativement bien selon les auteurs pour des lacs eutrophes et turbides dont les cyanobactéries sont en abondance. En fait, il est mentionné dans Simis *et al.* (2005) et dans Simis *et al.* (2006) que les prédictions des [PC] concordent mieux aux données de vérité terrain durant les périodes où les cyanobactéries dominent la biomasse du phytoplancton (rapport [PC]:[Chl] *a* élevé). Certaines limites ou erreurs par rapport à l'algorithme sont connues et répertoriées par Simis *et al.* (2005) : tout d'abord, l'absorption spécifique à 620 nm de la PC *in vivo* peut varier grandement dépendamment des conditions environnementales. Cette variabilité peut amener des erreurs dans les prédictions de l'algorithme. Aussi, selon la méthodologie de l'étude, l'absorption de la lumière à 620 nm devrait provenir uniquement de la PC après le retrait de l'apport de la Chl *a*. Par contre, il peut arriver, tel qu'indiqué par Simis *et al.* (2005), qu'une partie de l'absorption à 620 nm soit due aussi à d'autres pigments ou substances non considérés (autre que PC et Chl *a*). Cette manifestation est surtout importante dans le cas où la masse d'eau étudiée contient une grande variété de phytoplanctons sans être dominée par les algues bleu-vert. Cela a été prouvé par Simis *et al.* en 2006 lors d'une deuxième étude indépendante. Une campagne de terrain intensive a été menée aux Pays-Bas et en Espagne afin de valider le modèle semi-analytique déjà publié par Simis. Il a été trouvé que la Chl *b*, la phaeophytin et la fucoxanthin sont reliées à l'erreur de l'estimation de la [PC] par l'algorithme. Ces pigments pourraient influencer l'estimation du coefficient d'absorption à 620 nm, surtout pour des faibles concentrations en PC ou lorsque le rapport PC:Chl *a* est faible. Par contre, lorsque PC:Chl *a* est élevé, comme dans les lacs espagnols étudiés, le problème lié à l'absorption par d'autres pigments de phytoplanctons diminue (Simis *et al.*, 2006).

Enfin, la précision de cet algorithme a été validée pour les lacs eutrophes des Pays-Bas et d'Espagne ayant des concentrations en PC supérieures à 50 mg/m³ ce qui pourrait considérablement limiter la portée de l'algorithme si les manifestations de cyanobactéries sur les lacs du Québec ou du Canada n'ont pas des [PC] > 50 mg/m³. Par contre, si la proportion au Québec des différents pigments est sensiblement la même que ceux des lacs eutrophes de l'Espagne, alors la contribution de l'absorption de la Chl *b* (ou autre pigment secondaire) à 620 nm ne devrait pas être très importante et la précision de l'algorithme pourrait s'améliorer surtout

pour de faibles valeurs de [PC]. Dans tous les cas, l'application de l'algorithme sur des lacs du Québec nécessiterait sûrement une optimisation (une adaptation), car les caractéristiques optiques des deux lacs étudiés par Simis *et al.* (2005) diffèrent très probablement de celles rencontrées dans la province. De plus, il faudrait valider si les erreurs importantes en dessous de 50 mg/m^3 sont trop grandes pour permettre la mise en place d'un système de surveillance destiné à détecter les cyanobactéries avant, pendant et après des épisodes de fleurs d'eau. L'algorithme semi-empirique a été développé en utilisant les longueurs d'onde nominales du capteur MERIS de 300 m de résolution spatiale, ce qui est suffisamment précis pour l'étude de lac de "grandeur moyenne" selon Simis *et al.* (2005). Matthews *et al.* (2010) prétendent que le capteur MERIS a une résolution convenable pour les lacs d'une superficie supérieure à 1 km^2 . De nombreuses étendues d'eau du Québec ont donc une superficie assez grande telle que la Baie Missisquoi qui a une superficie d'environ 70 km^2 au total avec 50 km^2 sur le territoire québécois (OBVBM, 2011).

Le Tableau 3 ci-dessous est un rappel des trois algorithmes de la [PC] analysés dans ce mémoire.

Référence	Principe de l'algorithme
Dekker (1993)	Creux de réflectance à 624 nm par rapport à la ligne de base entre 600 et 648 nm.
Schalles et Yacobi (2000)	Rapport de bande, soit R_{650}/R_{625} .
Simis <i>et al.</i> (2005)	Algorithme semi-analytique qui emploie entre autres deux rapports de réflectance, soit R_{709}/R_{665} et R_{709}/R_{620} (compatible avec MERIS).

Tableau 3 : Liste des algorithmes de la [PC] présentés dans cette section.

La performance des trois algorithmes de télédétection de la PC mentionnés plus tôt (ceux de Dekker, Schalles et Yacobi, et Simis *et al.*) a été évaluée par Ruiz-Verdú *et al.* (2008). D'après leur étude, les erreurs les plus élevées proviennent de l'algorithme de rapport unique de réflectance (Schalles et Yacobi, 2000) alors que les plus faibles sont associées à l'algorithme semi-analytique (Simis *et al.*, 2005). Ruiz-Verdú *et al.* (2008) indiquent aussi que lorsque la [PC] $< 50 \text{ mg/m}^3$ (faible concentration mesurée), 54% des estimées pour l'algorithme de rapport unique de réflectance sont négatives, 31% le sont pour l'algorithme de ligne de base et enfin, 0,4% des estimations sont négatives pour l'algorithme semi-analytique. La performance varie encore une fois lorsqu'un sous-ensemble des données représentant les eaux fortement dominées par les cyanobactéries ($>80\%$ de la biomasse totale) est employé. En effet, Ruiz-Verdú *et al.* (2008) rapportent que les erreurs résiduelles ($[PC_{\text{estimés}}] - [PC_{\text{mesurés}}]$) de l'algorithme de rapport

unique de réflectance et de l'algorithme de la ligne de base augmentent pour ce sous-ensemble. Dans cette situation, seul l'algorithme de Simis *et al.* (2005) possède une régression significative par rapport aux mesures de [PC] même si l'algorithme a tendance à surestimer la [PC] pour les faibles concentrations et à la sous-estimer pour les fortes concentrations.

Malgré leurs approches et leurs résultats qui diffèrent, il a été démontré qu'en général les erreurs relatives de ces trois algorithmes sont plus élevées pour les faibles concentrations en cyanobactéries ($[PC] < 50\text{mg/m}^3$). De plus, une sous-estimation de la PC est observée de façon générale pour les concentrations élevées ($[PC] > 200\text{mg/m}^3$) [Ruiz-Verdú *et al.*, 2008]. Néanmoins, malgré certaines lacunes, l'algorithme de Simis *et al.* (2005) semble le plus précis des trois, car une correction de l'absorption à 620 nm causée par un autre pigment majeur (la Chl *a*) est prise en considération (*ibid.*). Notons aussi qu'il a été démontré par les résultats que la Chl *a* est le pigment qui influence le plus les prédictions des algorithmes (accentue les erreurs d'estimation) [*ibid.*].

3.3 Approche par variabilité de 2^e ordre

Les algorithmes de la littérature qui tentent d'estimer la [PC] ont ainsi certaines lacunes et l'exploration de nouvelles formulations pour la télédétection des cyanobactéries est une option à investiguer. Selon Brown *et al.* (2008), les algorithmes empiriques utilisés pour estimer la concentration en chlorophylle à partir de la réflectance pour les eaux océaniques décrivent bien la contribution de la Chl *a* sur le signal total. On peut ainsi modéliser un spectre moyen de réflectance pour une concentration de chlorophylle mesurée ou estimée par télédétection. Puisque la chlorophylle explique la majorité des variations du spectre de réflectance de l'océan, on parle ainsi de variabilité de 1^{er} ordre de la réflectance (même si la relation n'est pas mathématiquement de 1^{er} ordre nécessairement). Lorsque l'on soustrait ce spectre moyen du spectre mesuré, nous obtenons des résidus spectraux. Ces résidus ou anomalies contiennent de l'information sur d'autres variables affectant le signal (on parle alors de variabilité de 2^e ordre) et ils peuvent ainsi être utilisés pour étudier ces autres variables optiques. Nous croyons que l'application d'une démarche analogue dans les milieux lacustres pour détecter et quantifier l'influence des cyanobactéries sur la réflectance (anomalie causée par la PC) pourrait produire des résultats où l'approche de 2^e ordre est significativement mieux que l'approche de 1^{er} ordre.

4. Objectifs et hypothèses

4.1 Objectifs principal et secondaires

Objectif principal : Développer ou adapter une méthodologie permettant la surveillance des blooms de cyanobactéries sur les lacs du Québec à l'aide d'une approche empirique pour examinant les variabilités de 2^e ordre (anomalies) du spectre de réflectance (particulièrement celles causées par la PC) et comparer ses performances à celles d'autres algorithmes publiés.

Objectifs secondaires :

- Déterminer par régression deux relations empiriques entre des indicateurs de la présence de PC et un indicateur optique de la présence de biomasse, soit :

- (i) une relation entre la concentration de PC ([PC], mg/m³) et l'indicateur optique de la présence de biomasse phytoplanctonique (CI, sans unité). Cette relation exprime notre variabilité de 1^{er} ordre.

- (ii) une relation entre un indicateur optique de [PC] (PCI, sans unité) et CI.

- Déterminer une relation empirique entre les résidus (anomalies) des deux relations obtenues à l'objectif spécifique précédent. Cette relation obtenue décrira le lien entre les résidus de la [PC] ($\Delta[PC]$) et les résidus de PCI (ΔPCI). Elle utilise la variabilité de 2^e ordre des spectres de réflectance.

4.2 Hypothèses principale et secondaire

Hypothèse principale : Une amélioration de la surveillance des cyanobactéries est possible grâce à l'utilisation des variabilités de 2^e ordre du spectre de réflectance. Un lien empirique entre la concentration en PC et ces variabilités permet d'estimer la [PC] avec une meilleure précision que les algorithmes existants de la littérature.

Secondaire :

- La [PC] et PCI sont toutes deux corrélées avec la biomasse de phytoplancton (représentée ici par CI).

5. Méthodologie

5.1 Site à l'étude

Deux jeux de données distincts seront utilisés lors de ce projet de recherche. Un premier est fourni par Stefan Simis [utilisé dans Simis *et al* (2006)]. Ces derniers ont mené une campagne terrain intensive entre 2001 et 2005 sur différents lacs et réservoirs d'Espagne et des Pays-Bas (57 différents au total). Les spectres de réflectance *in situ* et les concentrations en Chl *a* et en PC ont été obtenus sur chaque étendue d'eau. Ceci constitue une base de données idéale pour le développement d'un algorithme de télédétection des cyanobactéries. Notons que les auteurs ont utilisé les spectromètres ASD-JR et PR-650 pour la prise de mesures spectrales [les spécifications techniques de ces spectroradiomètres sont décrites dans ASD (2009) et Photo Research (2010) respectivement]. De plus, deux méthodes pour l'extraction de la [PC] furent employées en laboratoire (cycle répété de gel et dégel et broyage mécanique) en plus d'une méthode de mesure *in situ* de la PC par fluorescence [voir Simis *et al* (2006) pour les détails].

Une autre série de mesures a été acquise durant l'été 2009 par notre équipe sur la baie Missisquoi, le lac Memphrémagog et le lac St-Pierre afin de vérifier l'applicabilité de l'algorithme au Québec. À l'été 2010, la baie Missisquoi a été visitée de nouveau pour compléter ce 2^e jeu de données. Les trois étendues d'eau ont été choisies en fonction de leur dimension, leur juridiction (fédérale) et de leur susceptibilité à contenir des cyanobactéries. Tout d'abord, elles ont une dimension suffisamment grande pour être étudiées par des capteurs satellitaires de résolutions spatiales d'environ 300 m comme MERIS. La responsabilité de la juridiction fédérale implique la surveillance des cours d'eau ou des lacs qui touchent une frontière provinciale ou internationale ou encore qu'ils soient un cours d'eau navigable (les trois étendues d'eau remplissent cette condition). Cet aspect est important à considérer, car les campagnes terrains sont faites en collaboration avec Environnement Canada. Enfin, des cyanobactéries ont déjà été répertoriées dans la baie Missisquoi (Lavoie *et al.*, 2007), le lac Memphrémagog (Lacasse Benoit, 2009) et le lac St-Pierre (des cyanobactéries benthiques, voir Centre Saint-Laurent, 2009).

Sur chacune des étendues d'eau, l'état de l'eau et les conditions environnementales varient spatialement. Ainsi, un soin particulier est consacré au positionnement des sites d'échantillonnage afin d'avoir un portrait global et représentatif de toutes les conditions présentes

sur les masses d'eau à l'étude. De plus, tous les sites sélectionnés sont situés à plus de 300 m des berges ce qui correspond à la résolution spatiale de MERIS. De cette façon, les stations étudiées ne correspondent pas à des pixels mixtes couvrant à la fois la terre et l'eau. Aussi, les effets de fonds sont minimisés en évitant l'eau des berges souvent peu profonde. Dix stations par plan d'eau sont choisies pour couvrir la baie Missisquoi et le lac Memphrémagog (Figure 4 et Figure 5 respectivement) alors que 22 stations sont retenues pour le lac Saint-Pierre en raison de sa grande dimension et de son hydrodynamisme plus complexe (Figure 6).

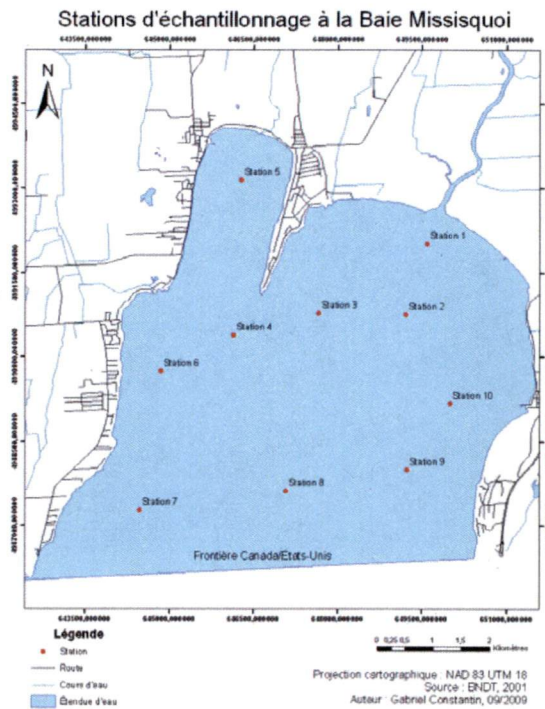


Figure 4 : Carte des stations d'échantillonnage à l'étude dans la baie Missisquoi.

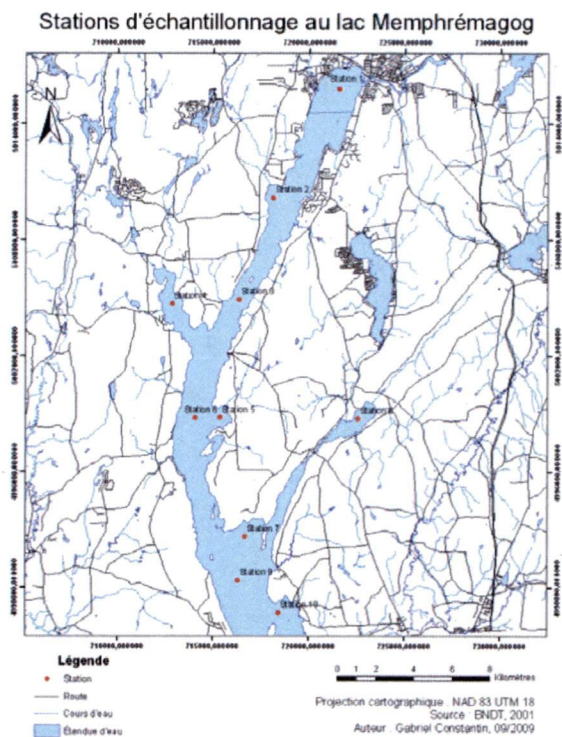


Figure 5 : Carte des stations d'échantillonnage à l'étude dans le lac Memphrémagog.

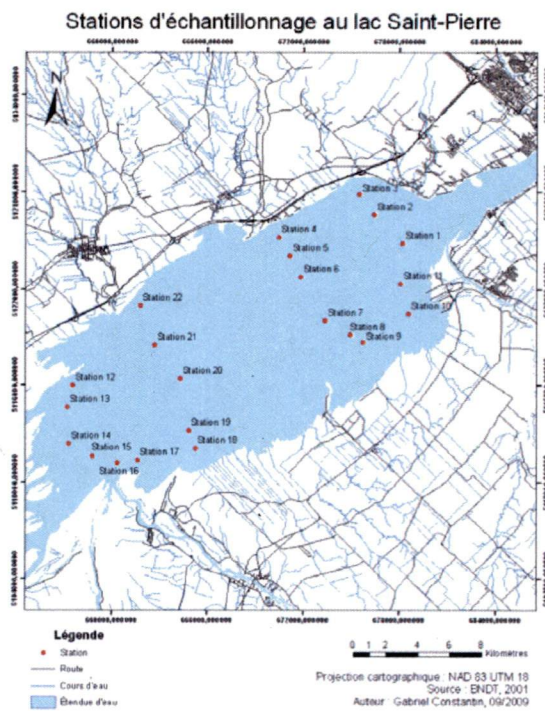


Figure 6 : Carte des stations d'échantillonnage à l'étude dans le lac Saint-Pierre.

5.2 Étapes méthodologiques

Dans le but de bien illustrer les étapes méthodologiques du projet de recherche, un organigramme est présenté à la Figure 7.

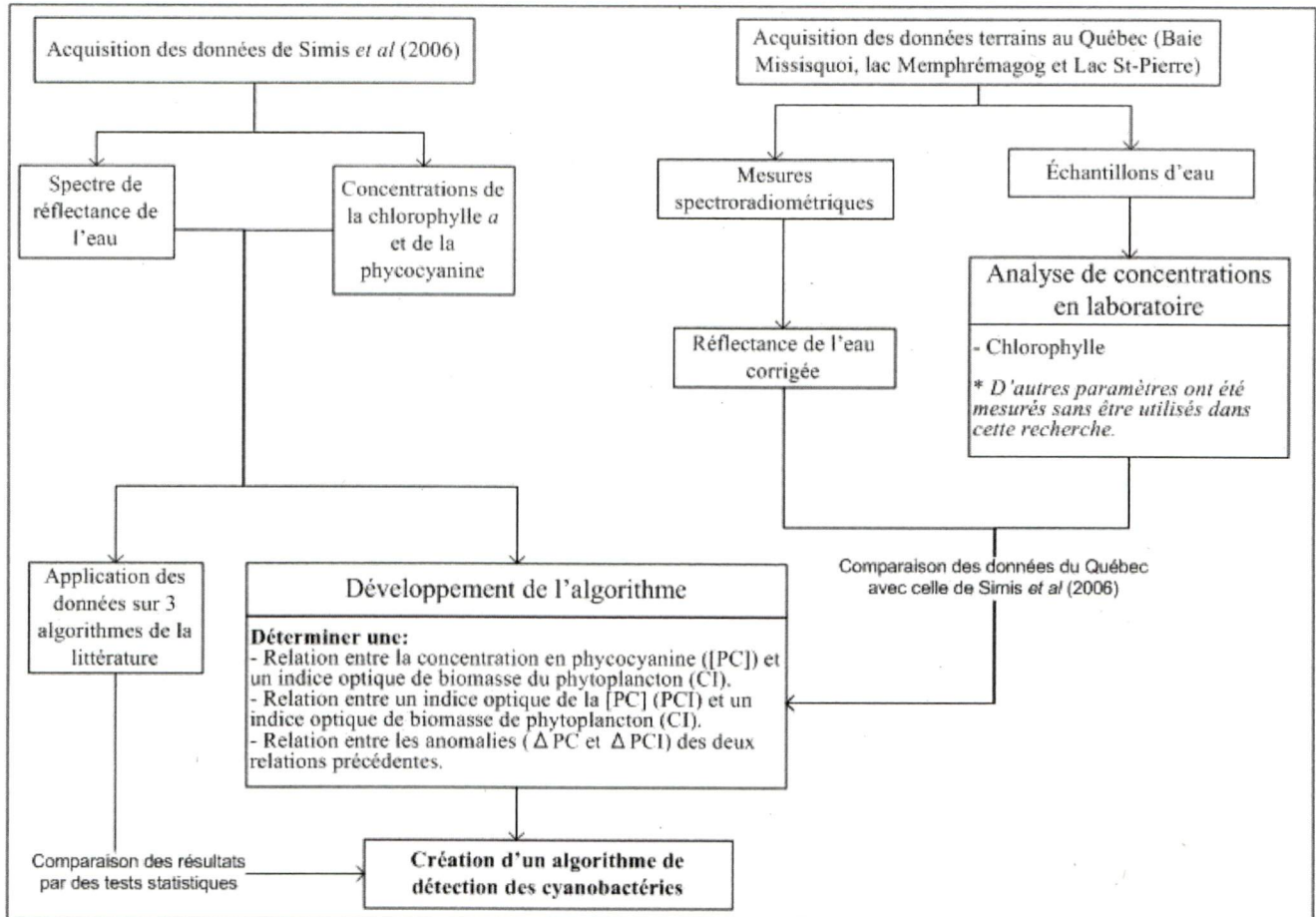


Figure 7 : Organigramme méthodologique

Mesures spectroradiométriques

Les mesures de réflectance de l'eau (ρ_{eau}) sont prises au Québec (campagne 2009 et 2010) avec le spectromètre ASD-Junior (angle de visée de 25° et domaine spectral de 350-2500 nm). Un Spectralon 5% retournant une luminance noté $L_{spect5\%}$ est surface plate qui reflète 5% de l'éclairement descendant et pour laquelle la distribution angulaire de la luminance réfléchi est approximativement isotropique ou Lambertienne. Elle est nécessaire pour calculer l'éclairement descendant (E_d , $W m^{-2} nm^{-1}$), car cette dernière est employée pour éviter les grandes différences de signal entre la luminance ascendante provenant de l'eau (L_{eau} , $W m^{-2} nm^{-1} steradian^{-1}$) et la

luminance provenant d'un spectralon de 100% (l'eau réfléchit ou rétrodiffuse une faible fraction de l' E_d et donc l'éclairement qu'elle renvoie est beaucoup plus faible que E_d). Le traitement effectué sur les spectres de compte numérique (CN) est le suivant pour un cas typique d'une mesure de l'eau. La mesure de l'eau en compte numérique (CN_{eau}) est normalisée par la mesure du Spectralon 5% ($CN_{spect5\%}$) :

$$\frac{CN_{eau} - DK}{20(CN_{spect5\%} - DK)} = \frac{L_{eau} / k}{20(L_{Ed} / k)} = \frac{L_{eau}}{20 L_{spect5\%}} \quad (3)$$

où DK = Dark current, $L_{spect5\%}$ = Luminance réfléchi par le Spectralon de 5% ($W m^{-2} nm^{-1} steradian^{-1}$) et k = constante de calibration entre la luminance et les comptes numériques de l'ASD. Un facteur de 20 (soit $1 / 0.05$) est appliqué au dénominateur pour ramener les conditions d'éclairement à 100%. Tous les paramètres ci-haut sont dépendants de la longueur d'onde (λ), on n'utilise aucun argument explicite en λ afin de simplifier les équations. Puisque la surface du Spectralon est considérée approximativement comme une surface Lambertienne, alors on peut démontrer que l'éclairement descendant est $E_d = 20 L_{spect5\%} \pi$ (Ryer, 1997). Ainsi, en supposant une émission isotropique de l'eau, on peut affirmer que :

$$\frac{L_{eau}}{20 L_{spect5\%}} = \frac{L_{eau}}{\frac{E_d}{\pi}} = \pi \frac{L_{eau}}{E_d} = \rho_{eau} \quad (4)$$

Il faut noter que les réflectances obtenues sont en fait des réflectances de télédétection ("remote sensing reflectances") puisque l'éclairement ascendant (E_u) est substitué par πL_{eau} (une substitution logique puisque les capteurs de télédétection peuvent seulement mesurer la luminance). La luminance du ciel normalisée par la luminance du Spectralon 5% (proportionnelle à l'éclairement descendant) est aussi acquise à chacune des stations. Cette mesure est prise afin de corriger l'apport de la réflexion spéculaire du ciel sur les spectres de réflectance de l'eau. Le processus de correction du ciel appliqué aux spectres de réflectance de l'eau acquis au Québec est décrit à l'annexe 2. Puis, l'annexe 3 présente la démonstration physique de la formule utilisée pour corriger la réflexion du ciel. Enfin, pour connaître les définitions fondamentales des paramètres optiques ci-dessus, il suffit de consulter l'ouvrage de Mobley (1994).

Le protocole terrain pour la prise de mesures spectrales suit les recommandations de Mueller *et al* (2003) (angle de visée respectable, positionnement sur le bateau, techniques pour éviter les réflexions spéculaires, etc.). Tout d'abord, les ombres et les réflexions spéculaires sont évitées en plaçant le viseur et le Spectralon loin de toute surface qui pourrait réfléchir la lumière sur ce dernier (Mueller *et al*, 2003). Dans le même ordre d'idées, il est préférable de se positionner soit à l'avant ou au coin arrière du bateau pour la prise de mesure de l'eau, car nous évitons ainsi le plus possible les réflexions spéculaires de la lumière sur la coque du bateau (réduis la surface de réflexion effective face à la mesure optique). Il est important aussi d'esquiver les vagues et/ou la mousse provoquée par le mouvement du bateau (*ibid.*). Pour minimiser les réflexions du ciel, du soleil et du bateau à la surface, notons que pour chaque mesure de l'eau ou du ciel, l'angle zénithal est de 45° par rapport au nadir et l'angle azimutal est de 135° par rapport au soleil (*ibid.*). Vingt spectres sont acquis pour chaque cible étudiée. La moyenne de ces derniers permet d'éliminer une partie du bruit aléatoire généralement présent quand une seule mesure est prise (ASD, 1999). Le risque d'obtenir qu'une seule mesure qui proviendrait du scintillement du soleil sur la surface de l'eau en raison des vagues est minimisé aussi. Il faut noter que les mesures spectrales et les échantillons d'eau sont pris simultanément et au même endroit (obtention de paramètres comparables). Enfin, des critères de qualité sont établis pour les spectres de réflectance afin d'éliminer les spectres provenant d'une mauvaise mesure.

Mesures des paramètres de qualité de l'eau

Lors des campagnes de terrains au Québec, plusieurs échantillons d'eau sont récoltés et analysés en laboratoire afin d'obtenir une série de paramètres de qualité de l'eau. Une série des paramètres physico-chimiques de qualité de l'eau est obtenue près de la surface grâce à une multisonde YSI 6600 V2-4 qui est immergée dans l'eau. Tous les paramètres suivants sont acquis directement sur le terrain : la température de l'eau, la conductivité, le pH, la turbidité, l'oxygène dissous, la fluorescence, la concentration en chlorophylle et en phycocyanine par fluorimétrie. De plus, des échantillons sont pris pour la concentration en chlorophylle, en matière en suspension, en carbone organique dissous et total, en phosphore total, en nitrite-nitrate et en cyanobactérie (intervalles du nombre de cellules [exemple : entre 10 000 et 20 000] des espèces potentiellement toxiques et pour 6 échantillons, décompte et identification complets de toutes les espèces d'algues). Ils sont analysés par la suite en laboratoire. Pour les paramètres de qualité de l'eau cités ci-dessus,

l'American Public Health Association (2005) décrit bien ce que représente chacun d'eux et comment ils sont obtenus en laboratoire. Selon les besoins de la présente étude, la [Chl *a*] est le paramètre obtenu au Québec le plus pertinent. Des analyses du nombre de cellules de cyanobactéries ont été effectuées par le Centre d'expertise en analyse environnementale du Québec. Les résultats renvoyés ne sont pas convenables pour le développement d'un algorithme (intervalles du nombre de cellules uniquement) et aucune mesure de [PC] n'a été réalisée. Par contre, les données du Québec serviront à vérifier si l'indice de biomasse s'applique aussi bien aux sites d'étude de la province que ceux d'Europe. En ce qui a trait au développement de notre algorithme, ce sont donc les mesures de [PC], de [Chl *a*] et les spectres de réflectance obtenus par Simis *et al* (2006) qui sont employées (les analyses en laboratoire et mesures terrains sont décrites dans leur étude).

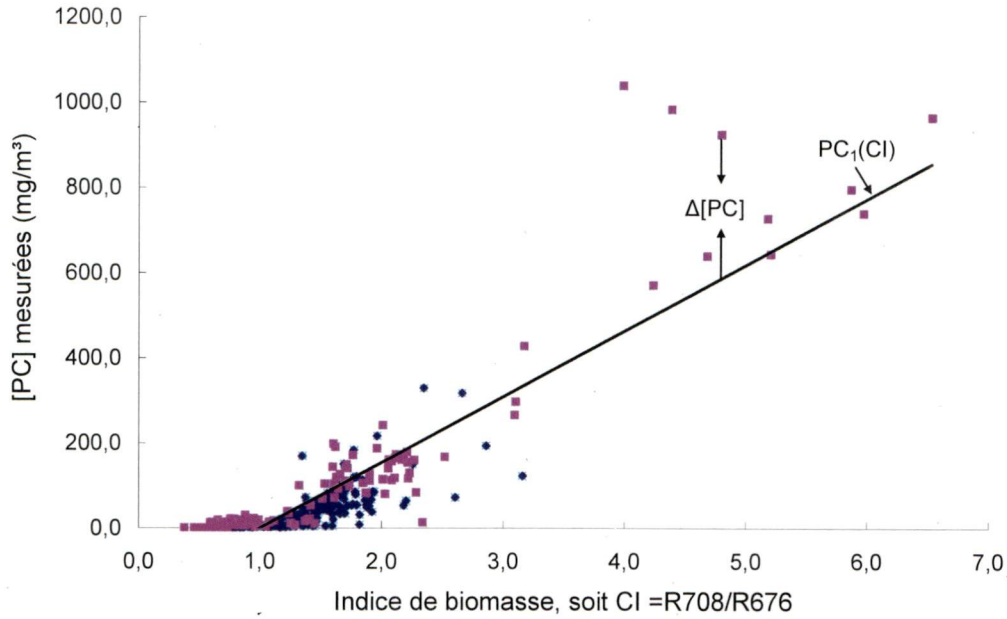
Développement de l'algorithme

Suite à l'acquisition des spectres de réflectance et des paramètres de qualité de l'eau, il est possible de débiter les analyses statistiques dans le but de retrouver la variabilité de 1^{er} ordre des spectres et les anomalies (variabilités de 2^e ordre) pertinentes au présent projet de recherche. Nous considérons ici que dans les lacs, la composante optique qui est associée avec le maximum de variance du spectre de réflectance dans les bandes allant de l'orange au rouge est la Chl *a* (variabilité de 1^{er} ordre). Tout d'abord, une relation est générée entre la concentration mesurée de cyanobactéries (exprimée par la [PC]) et un indice de biomasse du phytoplancton (soit CI pour *chlorophyll index*). CI est obtenu à partir du rapport entre la réflectance à 708 nm et celle à 676 nm (R708/R676, sans unité) qui est un bon indicateur de [Chl *a*] (voir Tableau 1). La droite obtenue par régression entre la [PC] mesurée et CI (soit $PC_1(CI)$, mg/mg³) représente la relation de 1^{er} ordre entre la quantité de cyanobactéries et la quantité totale de phytoplanctons (mesuré par leur impact sur la réflectance). L'algorithme qui tente d'estimer [PC] peut s'écrire sous la forme suivante :

$$[PC] = PC_1(CI) + \Delta[PC] \quad (5)$$

où CI est défini ci-dessus et $\Delta[PC]$ sont les résidus entre la [PC] et la droite $PC_1(CI)$. Les résidus ($\Delta[PC]$) représentent en fait les anomalies ou la variabilité de 2^e ordre de l'équation 5 (voir Figure

8 pour représentation graphique). Ils contiennent l'information additionnelle pour les cas où la quantité de cyanobactéries est significativement différente de celle qui est estimée par la formule de régression sur la Figure 8. La Figure 8 présente la relation entre [PC] et CI générée à partir des données venant de Simis *et al* (2006). Le graphique illustre également la droite de régression $PC_1(CI)$ et les anomalies de la relation ($\Delta[PC]$).



**Figure 8 : Relation de 1^{er} ordre entre les [PC] mesurées et l'indice de biomasse (CI).
Données des Pays-Bas en bleu et ceux de l'Espagne en violet.**

Pour utiliser l'équation 5 avec un algorithme de télédétection, il faut trouver une façon d'estimer $\Delta[PC]$ de façon optique. En ce sens, CI est ensuite comparé à un indice optique de la PC (PCI) qui est égal au rapport entre la réflectance à 650 nm et à 620 nm (R_{650}/R_{620} , sans unité). Ce rapport de bandes tiré de Schalles et Yacobi (2000) a un bon potentiel d'estimer adéquatement la quantité de [PC] grâce à l'utilisation de la bande à 620 nm. Cette nouvelle relation entre PCI et CI représente l'équivalent optique de l'équation 5. Elle peut s'exprimer ainsi :

$$PCI = PCI_1(CI) + \Delta PCI \quad (6)$$

La droite de régression $PCI_1(CI)$ retrouvée entre PCI et CI est soustraite des mesures (soit les valeurs de PCI) pour obtenir les anomalies en terme de réflectance, soit ΔPCI de l'équation 6.

Ainsi, deux types d'anomalies sont générés provenant des équations 5 et 6 ($\Delta[PC]$ et ΔPCI). Ces anomalies peuvent être comparées entre elles et une fonction les reliant est établie :

$$\Delta[PC] = \Delta PC_1(\Delta PCI) \quad (7)$$

où $\Delta PC_1(\Delta PCI)$ représente la relation empirique établie à partir d'une régression entre les quantités $\Delta[PC]$ et ΔPCI .

Un algorithme qui estime la $[PC]$ est développé désormais à partir des variabilités de 1^{er} et de 2^e ordre des spectres de réflectance de l'eau. L'indice de biomasse de phytoplancton (toujours appelé CI) permet dans un premier temps de donner une première estimation de la $[PC]$ présente (à partir de la fonction linéaire $PC_1(CI)$). Puis, l'ajout d'une correction de la $[PC]$ provenant de la relation entre les anomalies vient raffiner l'algorithme ($\Delta[PC]$ de l'équation 6). Ainsi, l'algorithme de télédétection qui permet d'estimer la $[PC]$ est donné par l'équation 5.

Application de l'algorithme de concentration de la phycocyanine

Lorsque de nouveaux spectres de l'eau sont acquis (des ensembles de mesures indépendantes), CI et ΔPC peuvent être retrouvés directement à partir des relations préalablement obtenues. Le processus d'estimation de la $[PC]$ pour chaque mesure (pour chaque spectre de réflectance) sera de :

- (i) calculer CI et PCI à partir des spectres mesurés.
- (ii) Établir les fonctions $PC_1(CI)$ et $PCI_1(CI)$ à partir des diagrammes de dispersion de $[PC]$ versus CI et de $[PC]$ versus PCI où CI et PCI parvient de l'étape (i).
- (iii) Calculer $\Delta(PCI)$ de l'équation 6 à partir de PCI de l'étape (i) et de $PCI_1(CI)$ de l'étape (ii) :

$$\Delta PCI = PCI - PCI_1(CI).$$
- (iv) Insérer les valeurs de $\Delta(PCI)$ dans l'équation 7 afin de calculer $\Delta[PC]$.
- (v) Insérer la valeur de $PC_1(CI)$ ainsi que $\Delta[PC]$ de l'étape (iv) dans l'équation 5 afin d'estimer la $[PC]$.

Comparaison avec d'autres algorithmes

Au moins trois algorithmes ont déjà été développés dans le but d'estimer la [PC] (Ruiz-Verdú *et al*, 2008). Il s'agit des algorithmes de Simis *et al* (2005), Schalles et Yacobi (2000) et de Dekker (1993). Les mêmes données servant à générer notre algorithme sont appliquées sur ces trois algorithmes. Une comparaison de la performance des algorithmes est effectuée entre autres grâce à des tests statistiques (par exemple le coefficient de détermination et RMSE).

6. Présentation, analyse et interprétation des résultats

6.1 Algorithme de variabilité de 2^e ordre

Tous les résultats de ce document qui servent directement à la vérification de notre algorithme d'estimation de la [PC] sont générés à partir d'un jeu de données acquis par Simis *et al* (2006) sur certains lacs et réservoirs des Pays-Bas et d'Espagne. En ce qui a trait aux données du Québec, uniquement les mesures spectroradiométriques et les mesures de [Chl *a*] sont disponibles. Cette contrainte nous limite grandement pour le développement d'algorithmes d'estimation de la [PC]. La Figure 9 illustre la relation entre le logarithme de [PC] et le logarithme de [Chl *a*] pour les échantillons des Pays-Bas et d'Espagne. Pour cette étude, le logarithme des variables a été adopté, car il nous permet d'obtenir une variance plus homogène pour l'ensemble des données. En effet, en prenant un logarithme des variables comme la [PC], la [Chl *a*] et CI, les anomalies générées à partir des relations de 1^{er} ordre sont réparties plus également autour de la droite moyenne. La relation entre $\text{Log}([PC])$ et $\text{Log}([Chl\ a])$ diffère selon le pays d'origine, soit les Pays-Bas ou l'Espagne (Figure 9). Ce phénomène est probablement dû aux espèces et aux conditions environnementales différentes. Il est donc difficile de lier directement la PC à la Chl *a*.

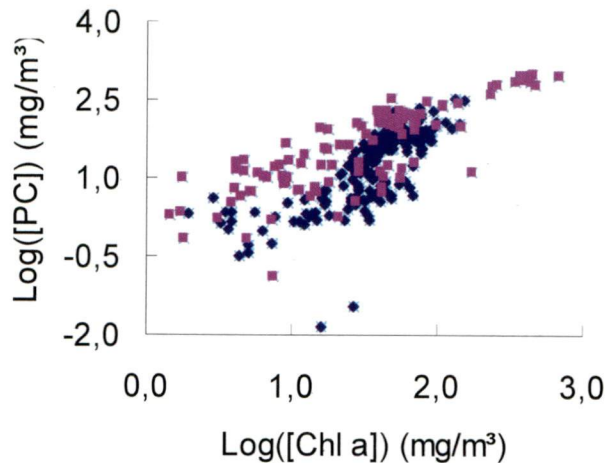


Figure 9 : Relation entre $\text{Log}([PC])$ mesurés et $\text{Log}([Chl\ a])$ mesurés. Résultats des Pays-Bas en bleu et ceux de l'Espagne en violet. $R^2 = 0,53$.

La Figure 10 décrit cette fois-ci la relation entre le rapport des concentrations des pigments et la concentration en Chl *a*, soit la relation entre $\text{Log}([PC]/[Chl\ a])$ et $\text{Log}([Chl\ a])$. Puisqu'il n'y a pas de corrélation significative entre ces deux variables pour la base de données complète (les données des Pays-Bas et de l'Espagne), on peut conclure qu'il n'y a pas de dépendance directe entre le rapport des deux pigments à l'étude et [Chl *a*]. C'est-à-dire que la fraction de la biomasse totale (en termes de pigment) formée par les cyanobactéries n'est pas une fonction de la biomasse

totale du phytoplancton. Par contre, pour les données des Pays-Bas uniquement, il semble y avoir une légère corrélation ($R^2=0,16$) entre le rapport la phycocyanine et la concentration en Chl *a* indiquant que les cyanobactéries deviennent une plus grande fraction de la biomasse totale lorsque cette dernière augmente. De plus, la relation entre $\text{Log}([\text{Chl } a])$ et $\text{Log}(\text{CI})$ suggère que l'indice optique CI ($R708/R676$) est suffisamment universelle pour toutes les régions géographiques de l'étude (Figure 11). Les données du Québec varient sensiblement de la même manière que celles acquises par Simis *et al* (2006). Donc, l'utilisation de CI comme indicateur de biomasse du phytoplancton au Québec n'est pas différent des autres endroits étudiés. Enfin, la Figure 12 illustre la corrélation entre $\text{Log}([\text{PC}])$ mesurées et $\text{Log}(\text{PCI})$ qui représente l'indice optique de [PC] utilisé dans cette étude.

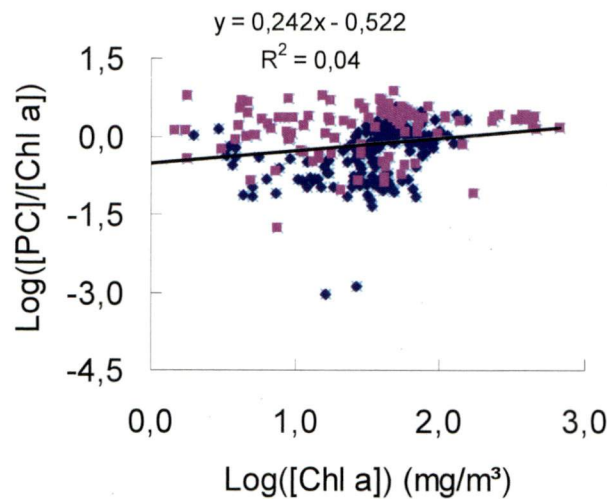


Figure 10 : Relation entre $\text{Log}([\text{PC}]/[\text{Chl } a])$ et $\text{Log}([\text{Chl } a])$ mesurés. Résultats des Pays-Bas en bleu et ceux de l'Espagne en violet.

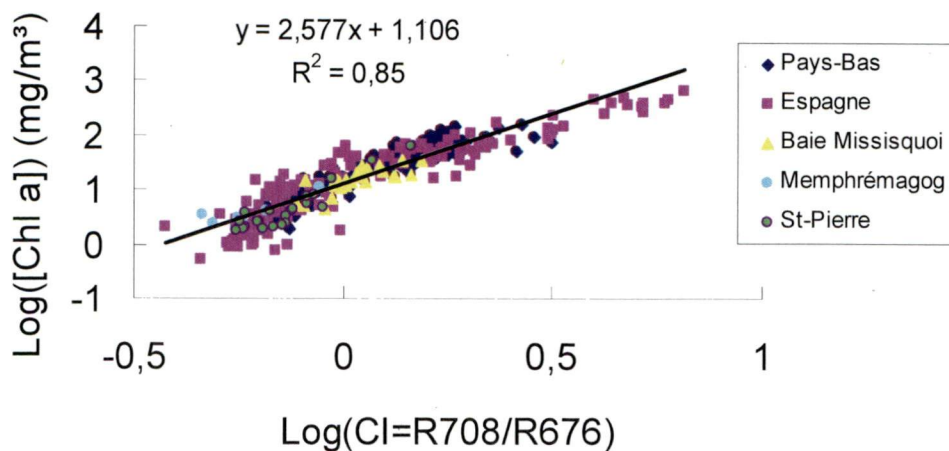


Figure 11 : Relation entre $\text{Log}([\text{Chl } a])$ mesurés et $\text{Log}(\text{CI}=\text{R708}/\text{R676})$ pour le jeu de données de Simis (Pays-Bas et Espagne) et les données acquises au Québec dans la baie Missisquoi, le lac Memphrémagog et le lac Saint-Pierre.

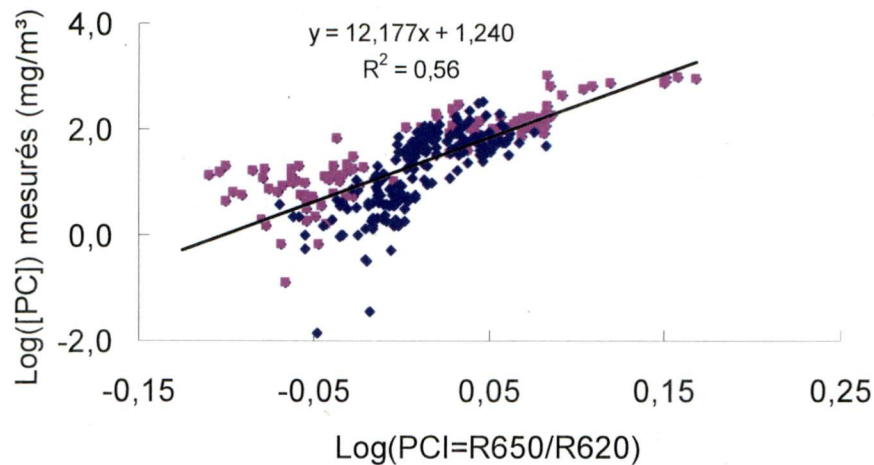


Figure 12 : Relation entre $\text{Log}([\text{PC}])$ mesurés et $\text{Log}(\text{PCI}=\text{R650}/\text{R620})$. Résultats des Pays-Bas en bleu et ceux de l'Espagne en violet.

Les résultats de l'algorithme initialement proposé dans ce projet de maîtrise sont présentés dans les prochains paragraphes. La droite de régression de la Figure 13 correspond à la relation de 1^{er} ordre analogue à l'équation 5, soit le terme $\text{PC}_1(\text{LogCI})$. C'est à partir de celle-ci que les résidus $\Delta\text{Log}([\text{PC}])$ sont calculés. La Figure 14 illustre cette fois-ci la relation entre PCI et CI et la droite $\text{PCI}_1(\text{CI})$ représente la relation de 1^{er} ordre de l'équation 6. Ainsi, les deux relations de 1^{er} ordre (droites des Figure 13 et Figure 14) permettent de calculer respectivement les anomalies entre $\text{Log}([\text{PC}])$ et $\text{Log}(\text{CI})$ et entre PCI et CI ($R^2 = 0,66$ pour les deux relations). Ces résultats nous confirment que $\text{Log}([\text{Chl } a])$ est une fonction modérément bonne de $\text{Log}([\text{PC}])$ (voir aussi la Figure 9) et que l'indice CI est un assez bon estimateur de la $[\text{Chl } a]$ (voir aussi Figure 11) [la variation de la concentration de Chl a explique les deux-tiers ($R^2=0.66$) de la variation en PC].

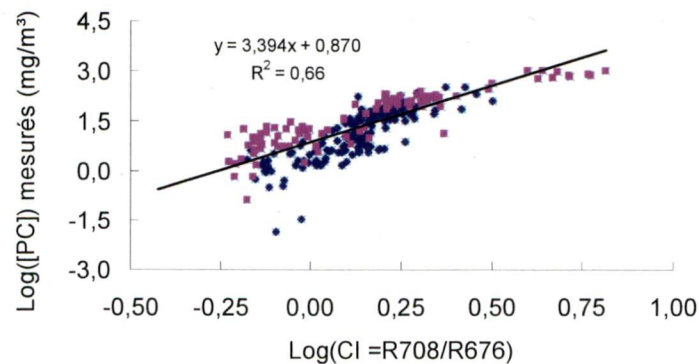


Figure 13 : Relation de 1^{er} ordre entre $\text{Log}([\text{PC}])$ mesurés et $\text{Log}(\text{CI}=\text{R708}/\text{R676})$. Résultats des Pays-Bas en bleu et ceux de l'Espagne en violet.

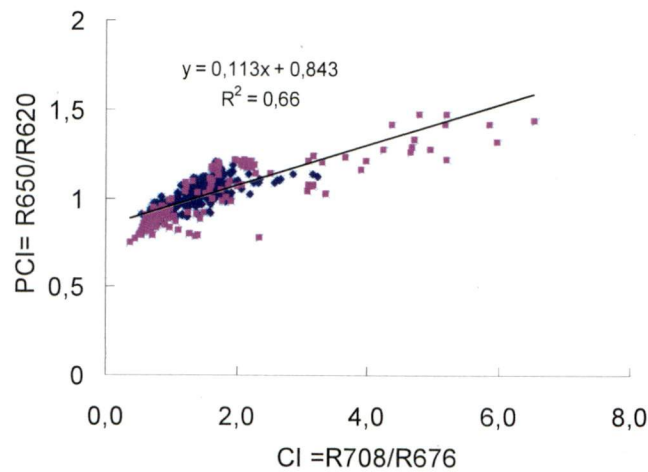


Figure 14 : Relation de 1^{er} ordre entre PCI et CI. Résultats des Pays-Bas en bleu et ceux de l'Espagne en violet.

La relation de 2^e ordre (Figure 15) est générée à partir des anomalies des deux premières relations (Figure 13 et Figure 14). Cette relation est incluse dans l'algorithme de [PC] afin d'enrichir l'information spectrale fournie par la relation de 1^{er} ordre. La faible corrélation entre les deux paramètres indique que l'approche de 2^e ordre ne fonctionne pas très bien pour l'ensemble de nos données. Plusieurs tests ont été effectués dans le but de vérifier si $\Delta\text{Log}([\text{PC}])$ serait davantage corrélé avec des bandes ou rapports de bandes de réflectance, mais ce fut sans succès (voir l'Annexe 4 pour une description complète des tests).

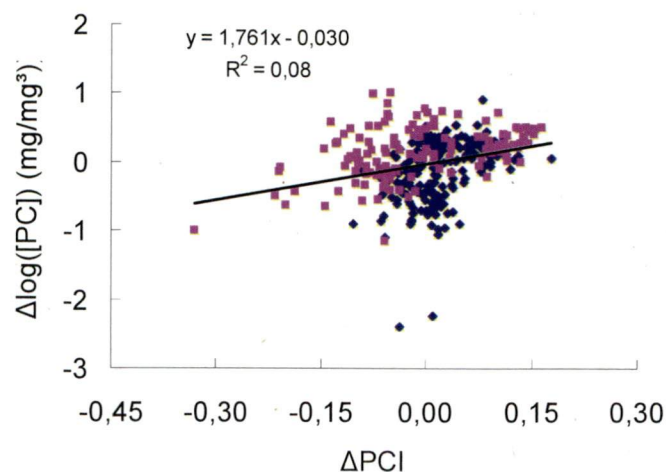


Figure 15 : Relation de 2^e ordre entre $\Delta\text{Log}([\text{PC}])$ ($\text{Log}([\text{PC}]) - \text{PC}_1(\text{LogCI})$) en fonction de ΔPCI ($\text{PCI} - \text{PCI}_1(\text{CI})$). Code de couleurs identique à la figure 2.

La Figure 16 ci-dessous est le résultat de l'ajout du terme de 2^e ordre obtenu à la Figure 15 pour améliorer l'indice CI de 1^{er} ordre et ainsi obtenir l'algorithme de [PC] corrigé. La relation entre les $\text{Log}([\text{PC}])$ prédits par notre algorithme de 2^{ième} ordre et les $\text{Log}([\text{PC}])$ mesurés (Figure 16a) semble meilleure pour les valeurs de $\text{Log}([\text{PC}])$ estimés au-dessus de $\sim 1,5$ ($R^2=0,69$ et $N=138$, Figure 16b), alors que les deux variables (les points des Pays-Bas et les points de l'Espagne) en dessous de ce seuil sont beaucoup moins bien corrélées ($R^2=0,27$ et $N=160$). La Figure 16b présente les mêmes résultats, mais uniquement pour les valeurs de $\text{Log}([\text{PC}])$ estimés au-dessus de 1,5 (au-dessus de la région de faible corrélation). La racine carrée de l'erreur quadratique moyenne (en $\log(\text{mg}/\text{m}^3)$), soit RMSE pour *root mean square error* dans l'espace logarithmique, est de 0,44 pour l'ensemble du jeu de données et de 0,25 pour le sous-ensemble de données comprenant uniquement les valeurs de $\text{Log}([\text{PC}])$ estimés au-dessus de 1,5. De plus, toujours pour le sous-ensemble de données, la pente est plus proche de 1 (0,926). Ces résultats démontrent que l'algorithme performe significativement mieux pour $[\text{PC}] > 10^{1,5} = 32 \text{ mg}/\text{m}^3$, car l'erreur relative décroît. Notons que 'N' correspond au nombre de données employées selon le cas étudié et l'approche linéaire précédente pourrait être remplacée par une régression multiple.

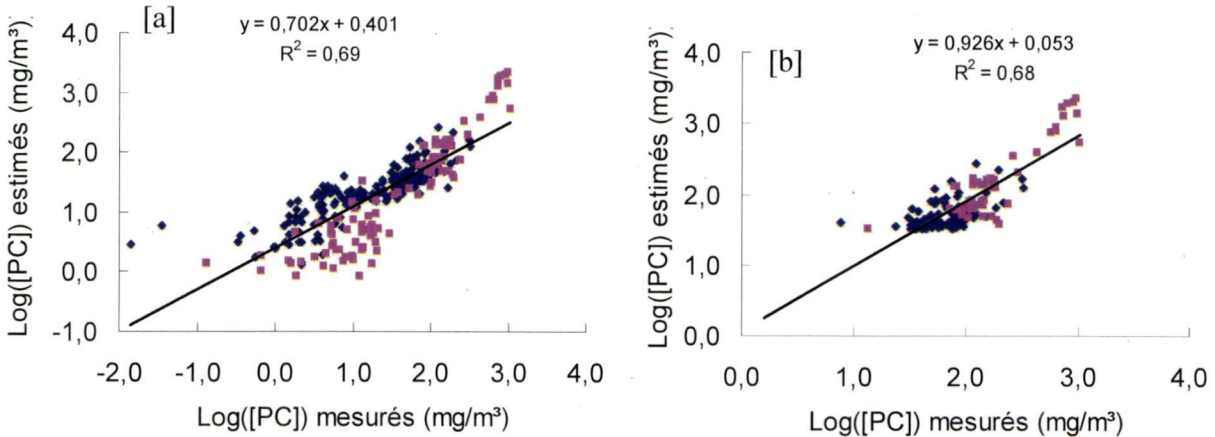


Figure 16 : Relations entre $\text{Log}([\text{PC}])$ estimés par notre algorithme de 2^{ième} ordre (équation 5) en fonction de $\text{Log}([\text{PC}])$ mesurés pour a) toutes les données et b) le sous-ensemble des données où $\text{Log}([\text{PC}])$ estimés est supérieur à 1,5. Résultats des Pays-Bas en bleu et ceux de l'Espagne en violet. RMSE=0,44 et 0,25 pour a) et b) respectivement.

Si nous employons uniquement la relation de 1^{er} ordre définie précédemment entre $\text{Log}([\text{PC}])$ et $\text{Log}(\text{CI})$, soit $\text{Log}([\text{PC}]) = \text{PC}_1(\text{Log}(\text{CI}))$, pour estimer $\text{Log}([\text{PC}])$, nous obtenons le résultat ci-dessous illustré à la Figure 17b ($R^2=0,50$ et $\text{RMSE}=0,35$). Il est possible d'observer que l'apport de la relation de 2^e ordre obtenu ci-dessus améliore significativement l'algorithme pour le sous-

ensemble restreint ($R^2=0,68$ et $RMSE=0,25$ de la Figure 16b) avec un seuil de confiance de 68% selon le test de signification retrouvé dans Siegel (1961) appliqué sur le coefficient de détermination. Par contre, il semble ne pas améliorer grandement l'algorithme pour le cas sans restrictions en [PC].

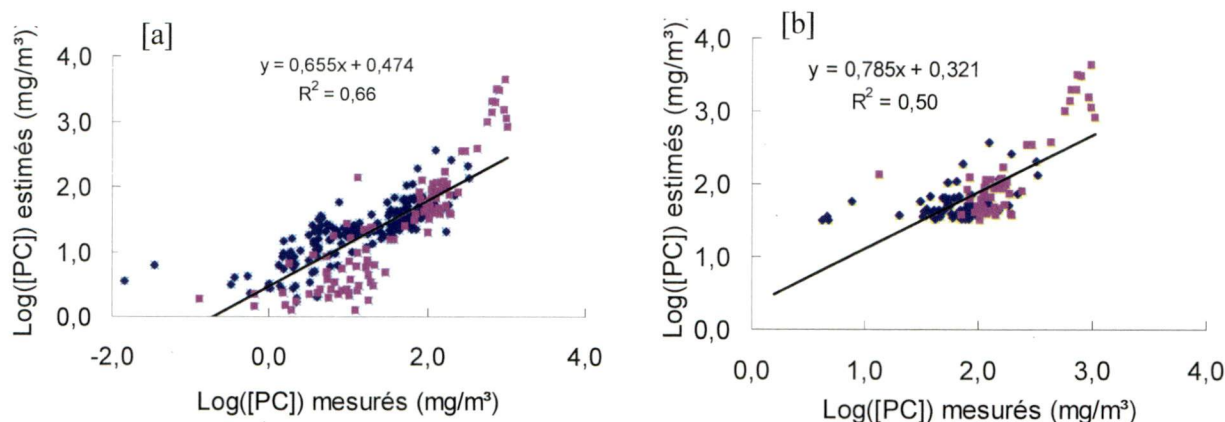


Figure 17 : Relation entre $\text{Log}([PC])$ estimés par la variabilité de 1^{er} ordre ($\text{Log}([PC])=PC_1(\text{LogCI})$) en fonction de $\text{Log}([PC])$ mesurés pour a) toutes les données et b) le sous-ensemble des données où $\text{Log}([PC])$ estimés est supérieur à 1,5. Résultats des Pays-Bas en bleu et ceux de l'Espagne en violet. $RMSE=0,46$ et $0,35$ pour a) et b) respectivement.

Nous avons testé aussi un algorithme de 2^e ordre semblable à l'exemple précédent qui emploie cette fois-ci le creux de réflectance à 620 nm calculé à partir de la ligne de base entre R568 et R650 (au lieu de $PCI = R650/R620$ de la Figure 14). Il a été possible de constater que l'algorithme de ligne de base de PCI qui calcule le creux de réflectance à 620 nm génère deux tendances très distinctes entre les données des Pays-Bas et celles d'Espagne. Il en résulte une estimation de $\text{Log}([PC])$ moins précise ($R^2=0,52$ et $RMSE=0,33$) par rapport à l'exemple précédent ($R^2=0,68$ et $RMSE=0,25$) pour les $\text{Log}([PC])$ estimés au-dessus de 1,5. Les algorithmes de ligne de base sont en fait influencés par d'autres variables comme les conditions de mesure qui peuvent modifier l'amplitude des spectres et fausser les résultats (Ruiz-Verdú *et al*, 2008). Par contre, on ne sait pas si ce problème est la cause de la performance apparemment inférieure de l'algorithme basé sur le creux à 620 nm.

Deux autres algorithmes de 2^e ordre quelque peu différent ont été analysés dans le but d'améliorer la performance et l'exactitude de l'algorithme de 2^e ordre initialement proposé. L'un des nouveaux algorithmes est développé à partir de la même démarche que notre algorithme de 2^e ordre initial, mais en employant cette fois-ci que les valeurs de $\text{Log}([PC])$ au-dessus de 1,5 dès la

première étape de création ($R^2=0,72$ et $RMSE=0,19$ pour $\text{Log}([PC])$ estimé au-dessus de 1,5). Puis, pour le deuxième algorithme alternatif de 2^e ordre, l'analyse des relations de 1^{er} et de 2^e ordre est faite séparément pour chacun des deux sites à l'étude, soit les Pays-Bas et l'Espagne ($R^2=0,78$ et $RMSE=0,36$ pour toutes les données et $R^2=0,63$ et $RMSE=0,24$ pour $\text{Log}([PC])$ estimés plus grand que 1,5). À première vue, ces résultats sont très intéressants au point de vue de leur performance. Une description plus détaillée de ses deux algorithmes est présentée à l'annexe 5. Dans la prochaine section, les résultats de différents algorithmes de télédétection de la [PC] provenant de la littérature seront présentés.

6.2 Comparaison avec d'autres algorithmes

La performance de différents algorithmes d'estimation de la [PC] tirés de la littérature est analysée dans cette section afin de les comparer entre eux et surtout avec l'algorithme de 2^e ordre développé dans cette étude. Notons que tous les résultats suivants sont générés à partir du même jeu de données que celui employé pour le développement de l'algorithme de 2^e ordre.

Tout d'abord, Simis *et al* (2005) emploient un algorithme semi-empirique relativement sophistiqué pour estimer la [PC]. Voici les équations 8, 9 et 10 que les auteurs utilisent, où a et b sont les coefficients d'absorption et de rétrodiffusion (m^{-1}) respectivement :

$$b(779)=1,61 \cdot R779 / [0,082 - 0,6 \cdot R779] \text{ [unités de } m^{-1}] \quad (8)$$

$$a(665)=147 \cdot [(R709/R620 \cdot (0,727 + b(779))) - b(779) - 0,401] \text{ [unités de } m^{-1}] \quad (9)$$

$$[PC]=170 \cdot \{[(R709/R620 \cdot (0,727 + b(779))) - b(779) - 0,281] - [0,24 \cdot a(665)]\} \text{ [unités de } mg/m^3] \quad (10)$$

La Figure 18a et la Figure 18b présentent les résultats de l'algorithme de Simis *et al* (2005). Il est possible de constater que cet algorithme n'améliore pas les estimations de $\text{Log}[PC]$ pour toutes les données ($R^2=0,58$ et $RMSE=0,56$) ou pour la région restreinte où $\text{Log}[PC]$ estimé est supérieur à 1,5 ($R^2=0,51$ et $RMSE=0,39$) lorsqu'on les compare aux estimations de notre algorithme de 2^e ordre (voir les Tableau 4 et Tableau 5 ci-dessous à la fin de la section Résultat pour un résumé des statistiques comparatives).

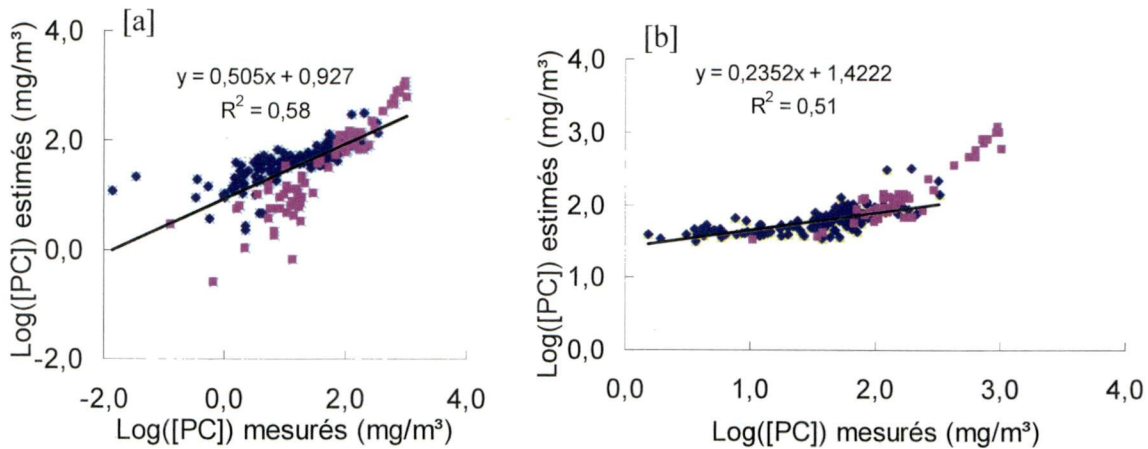


Figure 18 : Résultat de l'algorithme de Simis *et al* (2000) pour a) toutes les données et b) le sous-ensemble des données où Log([PC]) estimés est supérieur à 1,5. Résultats des Pays-Bas en bleu et ceux de l'Espagne en violet. RMSE = 0,56 et 0,39 pour a) et b) respectivement.

Nous pouvons observer que le rapport R709/R620 est employé dans l'équation 10 (ci-dessus) de l'algorithme de Simis *et al* (2005). La réflectance à 709 nm est influencée principalement par l'absorption de l'eau et la rétrodiffusion des cellules alors que la contribution causée par l'absorption des pigments à cette longueur d'onde est considérée comme étant négligeable. La réflectance à 620 nm, tel qu'indiqué auparavant, représente le maximum d'absorption de la PC. Afin de vérifier si la complexité de l'algorithme de Simis est nécessaire pour notre ensemble de données, nous avons développé notre propre algorithme simple et direct de 1^{er} ordre qui emploie uniquement le rapport de bande de l'équation 10 (Figure 19). La Figure 20 démontrent une bonne corrélation ($R^2=0,66$ et $RMSE=0,30$ pour le cas où Log([PC]) est supérieur à 1,5)) entre les prédictions de notre algorithme de rapport de bandes (R709/R620) et Log([PC]) mesurés. À la lumière de ces résultats et de ceux de l'algorithme de Simis *et al*. (2005), un simple algorithme empirique employant un rapport de bandes semble avoir un meilleur potentiel pour estimer la [PC] (coefficient de détermination significativement mieux selon le test de signification de Siegel [1961] avec un seuil de confiance de 68%). L'équation de l'algorithme est la suivante (R708 a été choisie au lieu de R709 en raison des mesures spectrales de notre jeu de données) :

$$\text{Log}([PC]) \text{ estimés} = 3,023 * \text{Log}(R708/R620) + 1,341 \quad (11)$$

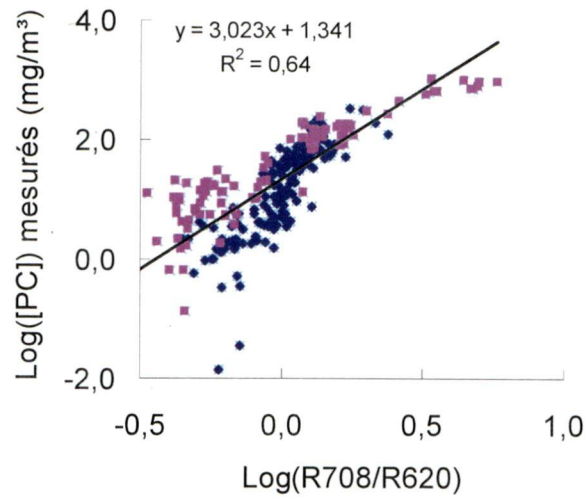


Figure 19 : Relation entre Log([PC]) mesurés en laboratoire en fonction du logarithme du rapport de bandes R708/R620. Résultats des Pays-Bas en bleu et ceux de l'Espagne en violet.

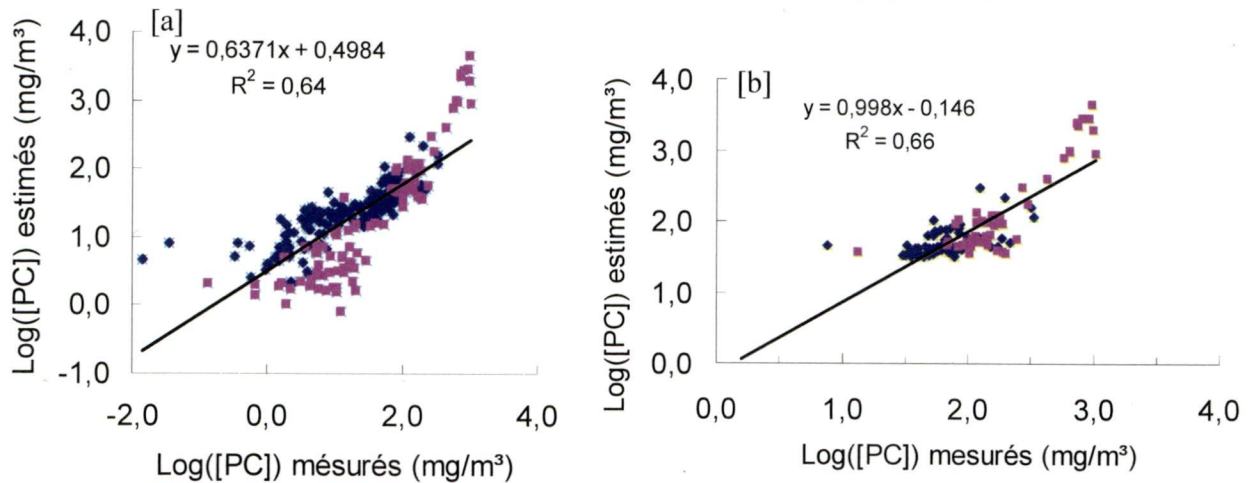


Figure 20 : Relation entre Log([PC]) estimés par l'algorithme de rapport de bandes (R708/R620) en fonction de Log([PC]) mesurés pour a) toutes les données et b) le sous-ensemble des données où Log([PC]) estimés est supérieur à 1,5. Résultats des Pays-Bas en bleu et ceux de l'Espagne en violet. RMSE = 0,47 et 0,30 pour a) et b) respectivement.

Deux autres algorithmes de la littérature qui tentent de retrouver la [PC] sont analysés. Les résultats de l'algorithme empirique de rapport de bandes par Schalles et Yacobi (2000) et de l'algorithme semi-empirique de ligne de base par Dekker (1993) sont illustrés aux Figure 21 et Figure 22 respectivement. L'équation de l'algorithme de Schalles et Yacobi (2000) est la suivante :

$$[PC] = (R650/R625 - 0,97) / 0,000912 \quad (12)$$

L'équation de l'algorithme de Dekker (1993) :

$$[PC] = -24,6 + 13686 * (0,5 * (R600 + R648) - R624) \quad (13)$$

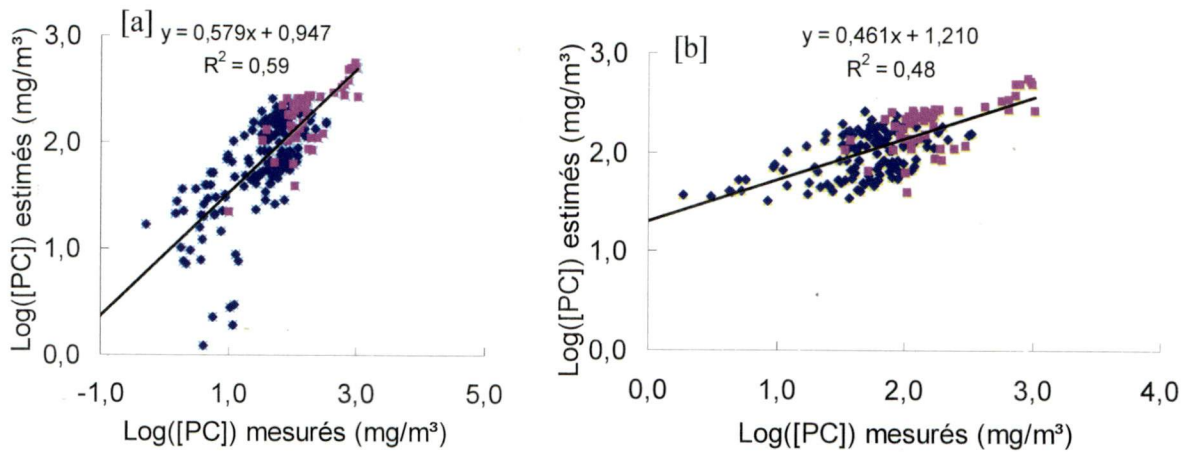


Figure 21 : Résultat de l'algorithme de Schalles et Yacobi (2000) pour a) toutes les données et b) le sous-ensemble des données où Log([PC]) estimé est supérieur à 1,5. Résultats des Pays-Bas en bleu et ceux de l'Espagne en violet. RMSE = 0,46 et 0,41 pour a) et b) respectivement

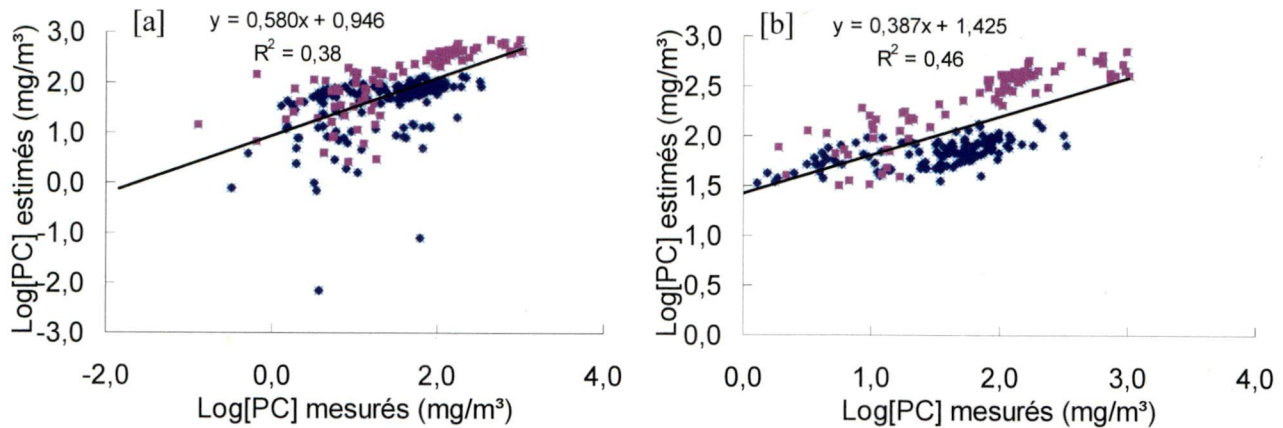


Figure 22 : Résultat de l'algorithme de Dekker (1993) pour a) toutes les données et b) le sous-ensemble des données où Log([PC]) estimé est supérieur à 1,5. Résultats des Pays-Bas en bleu et ceux de l'Espagne en violet. RMSE = 0,68 et 0,63 pour a) et b) respectivement.

Les Tableau 4 et Tableau 5 ci-dessous permettent de comparer la performance des différents algorithmes analysés dans cette étude, soit ceux provenant de la littérature et notre algorithme de variabilité de 2^e ordre. Les résultats contenus dans le Tableau 5 sont plus prometteurs qu'au Tableau 4, car les données en dessous du seuil critique de confiance ont été éliminées (c'est-à-dire la valeur de 1,5 du Log([PC]) estimé). En restreignant la valeur de Log([PC]) estimé à 1,5 ou plus, les résultats de tous les algorithmes sont améliorés selon le RMSE mis à part l'algorithme de Dekker (1993). En fait, cet algorithme de ligne de base est celui qui a les moins bons résultats suivi de l'algorithme de rapport de bande de Schalles et Yacobi (2000) selon les RMSE du Tableau 5. Par contre, cette constatation doit être nuancée, car un algorithme ayant un mauvais résultat au niveau du RMSE, mais un très fort R^2 pourrait être le meilleur des

algorithmes qui aurait simplement besoin d'être corrigé pour une erreur systématique (par exemple au niveau de l'extraction de la PC). L'algorithme de Simis *et al.* (2005) semble être le plus performant des trois provenant de la littérature. Cependant, il semble être moins bien adapté à l'estimation de la [PC] pour les plus faibles concentrations (Tableau 4). Notre algorithme de variabilité de 2^e ordre donne les meilleurs résultats d'après les analyses statistiques de RMSE autant lorsque toutes les données sont prises en compte ou bien pour le sous-ensemble de données. Pour le sous-ensemble de données où Log ([PC]) estimés est supérieur à 1,5, l'algorithme de 2^e ordre ($R^2=0,68$ et $RMSE=0,25$) a un coefficient de détermination significativement meilleur avec un seuil de confiance de 68% comparativement à chacun des trois algorithmes de la littérature à l'étude (sauf pour l'algorithme de Schalles et Yacobi optimisé, mais son RMSE est très élevé comparativement à celui de l'algorithme de 2^e ordre). Certaines variantes apportées à la méthodologie développée dans cette recherche permettent d'améliorer quelque peu la performance de notre algorithme de 2^e ordre initialement proposé. Entre autres, l'application aux données pour les Log([PC]) estimés supérieurs à 1,5 dès les premières étapes de la méthode génère le meilleur résultat observé ($RMSE=0,19$, voir Tableau 5). Puis, l'analyse de deux relations de 2^e ordre séparées selon les pays à l'étude permet d'augmenter la précision des estimations pour l'ensemble du jeu de données ($RMSE=0,36$, voir Tableau 4).

Nom ou référence de l'algorithme	PCI	Données d'entraînement	R^2	RMSE	N (n)
Var. de 2e ordre	R650/R620	Complet	0,69	0,44	298 (298)
Var. de 2e ordre	Creux de réflectance	Complet	0,66	0,44	298 (298)
Var. de 2e ordre séparée par pays (Esp. et P.-B.)	R650/R620	Complet	0,78	0,36	298 (298)
Var. de 2e ordre MERIS	R665/R620	Complet	0,68	0,44	298 (298)
Var. de 1 ^{er} ordre	-	Complet	0,67	0,46	298 (298)
Simis et al.	-	-	0,58	0,56	- (284)
Schalles et Yacobi	-	-	0,59	0,46	- (210)
Schalles et Yacobi après calibration	-	Complet	0,47	0,65	- (244)
Dekker	-	-	0,38	0,68	- (253)
Dekker après calibration	-	Complet	0,36	0,59	- (248)
Rapport R708/R620 tiré de Simis et al.	-	Complet	0,64	0,47	298 (298)

Tableau 4 : Statistiques des performances des différents algorithmes analysés dans cette étude pour toutes les données (Log([PC]) estimé versus Log([PC]) mesuré. PCI : l'indice optique de [PC] employé s'il y a lieu.

Données d'entraînement : indique si le jeu de données de Simis *et al.* (2006) a été employées. N (n) : "N" indique le nombre de données employées pour développer l'algorithme et "(n)" indique le nombre de données retournant des estimations positives de [PC].

Nom ou référence de l'algorithme	PCI	Données d'entraînement	R ²	RMSE	N (n)
Var. de 2e ordre	R650/R620	Complet	0,68	0,25	298 (138)
Var. de 2e ordre	Creux de réflectance	Complet	0,52	0,33	298 (131)
Var. de 2e ordre séparée par pays (Esp. et P.-B.)	R650/R620	Complet	0,63	0,24	298 (131)
Var. de 2e ordre	R650/R620	Log(PC)>1.5	0,72	0,19	136 (136)
Var. de 2e ordre MERIS	R665/R620	Complet	0,68	0,28	298 (130)
Var. de 1er ordre	-	Complet	0,50	0,35	298 (129)
Simis et al.	-	-	0,51	0,39	- (201)
Schalles et Yacobi	-	-	0,48	0,41	- (177)
Schalles et Yacobi après calibration	-	Complet	0,64	0,58	- (218)
Dekker	-	-	0,52	0,51	- (219)
Dekker après calibration	-	Complet	0,36	0,48	- (177)
Rapport R708/R620 tiré de Simis et al. (2005)	-	Complet	0,66	0,30	298 (137)

Tableau 5 : Statistiques des performances des différents algorithmes analysés dans cette étude pour le sous-ensemble des données où Log([PC]) estimés est supérieur à 1,5 ([PC] >32mg/m³). PCI : l'indice optique de [PC] employé s'il y a lieu. Données d'entraînement : indique si le jeu de données de Simis et al. (2006) a été employées. N (n) : "N" indique le nombre de données employées pour développer l'algorithme et "(n)" indique le nombre de données retournant des estimations positives pour les données dont le Log([PC]) est supérieur à 1,5.

En plus des tests statistiques, une analyse entre les erreurs relatives des estimations [Log([PC]) estimés - Log([PC]) mesurés] et le rapport des concentrations mesurées des pigments [Log([PC] / [Chl *a*])] vient renchérir nos observations au niveau de la performance des algorithmes. Le terme [Log([PC] / [Chl *a*])] démontre en fait la proportionnalité entre les deux pigments. Tous les résultats et figures décrivant cette analyse sont présentés à l'Annexe 6. Suite à ces derniers résultats, il est possible de conclure que les quatre algorithmes principaux à l'étude (algorithme de 2^e ordre et les trois de la littérature) génèrent des erreurs plus faibles lorsque le rapport Log ([PC]/[Chl *a*]) est élevé, soit lorsque les cyanobactéries deviennent de plus en plus dominantes à la surface de l'eau. Il a été démontré aussi auparavant que les algorithmes performant mieux pour les concentrations en [PC] plus élevées (Log[PC]>1,5 ou [PC]>32 mg m⁻³). Il serait désormais intéressant de vérifier quels algorithmes performant mieux pour les faibles concentrations en [PC]. La relation entre Log([PC]) estimés et Log([PC]) mesurés pour les valeurs de Log([PC]) estimés inférieures à 1,5 est présentée ci-dessous à la Figure 23 et Figure 24 respectivement pour notre algorithme et celui de Simis et al. (2005). Les figures pour l'algorithme de Dekker (1993) et de Schalles et Yacobi (2000) ne sont pas présentées, car elles ont un nombre plus limité d'estimations en dessous de 32 mg m⁻³. Il est possible de constater que notre algorithme de 2^e ordre est significativement meilleur que les autres algorithmes de la littérature pour les plus

faibles valeurs de [PC] avec un seuil de confiance de 95% selon le test de signification de Siegel (1961) pour le coefficient de détermination [voir Figure 23 et Figure 24 et le Tableau 6 décrivant les statistiques de cette dernière analyse]. Néanmoins, l'exactitude de l'algorithme de 2^e ordre n'est pas très grande ($R^2=0,27$ et $RMSE=0,55$) bien qu'elle soit considérablement supérieure à celle des autres.

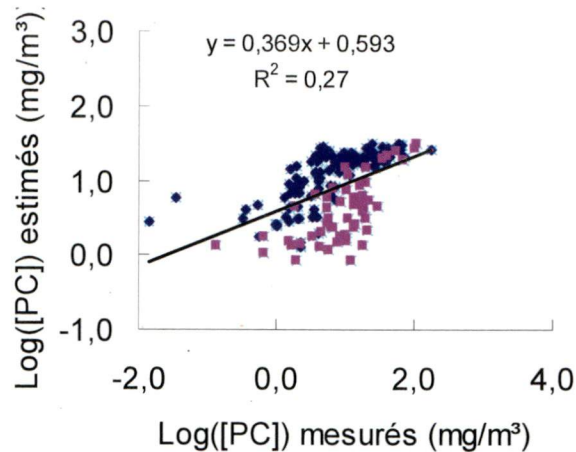


Figure 23 : Relation entre Log([PC]) estimés par l'algorithme de 2^e ordre et Log([PC]) mesurés pour le sous-ensemble de données où Log([PC]) estimés est inférieur à 1,5 ([PC] < 32mg/m³). $RMSE = 0,55$.

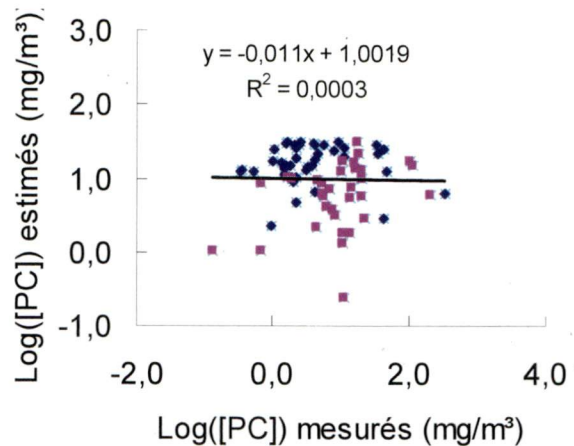


Figure 24 : Relation entre Log([PC]) estimés par l'algorithme de Simis *et al.* et Log([PC]) mesurés pour le sous-ensemble des données où Log([PC]) estimés est inférieur à 1,5 ([PC] < 32mg/m³). $RMSE = 0,81$

Nom ou référence de l'algorithme	R^2	RMSE	N
Var. de 2e ordre	0,27	0,55	160
Simis et al.	0,0003	0,81	78
Schalles et Yacobi	0,0028	0,67	33
Dekker	0,0055	0,83	60

Tableau 6 : Statistiques des performances des différents algorithmes analysés pour le sous-ensemble des données où Log([PC]) estimés est inférieur à 1,5 ([PC] < 32mg/m³). N : nombre d'estimations positives où Log([PC]) estimés est inférieur à 1,5.

7. Discussion des résultats

Selon les résultats présentés dans le Tableau 4 et le Tableau 5, l'algorithme de Dekker (1993) n'est pas en mesure d'estimer la [PC] pour notre ensemble de données ($R^2=0,38$ et $RMSE=0,68$). Cet algorithme de ligne de base a été développé pour des réflectances acquises sous l'eau et nous avons appliqué une transformation sur nos données prises au-dessus de la surface afin d'en assurer la compatibilité. En fait, les réflectances acquises au-dessus de l'eau ont été divisées par 0,543 pour les ramener à des conditions en dessous de la surface (Mueller *et al.*, 2003). Une telle transformation est importante étant donnée la nature de l'algorithme et malgré cet ajustement, la précision des estimations laisse à désirer. Les coefficients d'absorption et de rétrodiffusion ayant servi à développer l'algorithme semi-empirique de Dekker furent tirés de mesures prises sur 10 lacs eutrophes peu profonds dont la [PC] varie entre 7 et 113 mg/m^3 . Par contre, notre jeu de données provient de plus de 50 lacs et réservoirs ayant des conditions trophiques et des profondeurs différentes (Simis *et al.*, 2006) dont le 5^e centile est 0 mg/m^3 et le 95^e centile est 640,6 mg/m^3 pour les données d'Espagne (Ruiz-Verdú *et al.*, 2008). De ce fait, l'algorithme empirique de Dekker (1993) semble avoir été développé sous des conditions différentes. Une calibration de l'algorithme avec les données de qualité de l'eau de Simis *et al.* (2006) est donc une solution envisageable. Cette analyse supplémentaire a été faite et les résultats semblent avoir été améliorés au niveau du RMSE lorsque toutes les données sont considérées ($R^2=0,36$ et $RMSE=0,59$). Par contre, le coefficient de détermination s'abaisse considérablement pour l'algorithme calibré comparativement à l'algorithme initial de Dekker lorsque le sous-ensemble de $\text{Log}([PC])$ estimés au-dessus de 1,5 est analysé ($R^2=0,36$ et $RMSE=0,48$), alors que le RMSE reste sensiblement le même.

Quant à l'algorithme de rapport de bandes de Schalles et Yacobi (2000), il est possible de croire que la performance de cet algorithme ($R^2=0,48$ et $RMSE=0,41$ pour l'ensemble du jeu de données) serait supérieure si les auteurs avaient opté pour un autre rapport comme R708/R620 (plutôt que R650/R625) ou si on avait fait un ajustement des paramètres de l'algorithme pour les adapter à nos données. L'emploi de la bande de réflectance à 708 nm au lieu de celle à 650 nm a l'avantage que R708 a tendance d'augmenter proportionnellement avec la biomasse de phytoplancton et de compenser pour la diminution du coefficient d'absorption spécifique de la PC à 620 nm en raison du « *packaging-effect* » (Ruiz-Verdú *et al.*, 2008). En d'autres mots, la

réflectance à 708 nm (numérateur du rapport) augmente avec la biomasse ce qui compense l'effet de diminution moins rapide de la réflectance à 620 nm (dénominateur) en raison du « *packaging-effect* ». C'est ce phénomène qui permet au rapport R_{708}/R_{620} de mieux refléter la réalité malgré le « *packaging-effect* ». Néanmoins, l'algorithme de Schalles et Yacobi (2000) semble bien adapté à la forme des spectres de réflectances des Pays-Bas et d'Espagne. En effet, le pic du maximum de réflectance à 650 nm et le creux de réflectance à 625 nm employés dans l'algorithme de Schalles et Yacobi (2000) sont observables aux Figure 25 et Figure 26 qui illustrent respectivement la moyenne des spectres acquis au lac Carter du Nebraska (Schalles et Yacobi, 2000), et aux Pays-Bas et en Espagne. Bien que la forme concorde, les moyennes des valeurs de réflectance sont supérieures pour les spectres acquis sous l'eau par Schalles et Yacobi (2000) par rapport à nos spectres qui sont calculés juste au dessus de la surface. Cependant, Ruiz-Verdù *et al* (2008) ont démontré que l'amplitude des spectres n'a pratiquement aucune influence sur le résultat de cet algorithme (rapport de bandes) lorsque la forme des spectres est conservée, ce qui semble être le cas ici. De plus, ces mêmes auteurs prétendent que l'algorithme empirique peut être autant utilisé avec des réflectances sous l'eau ou au-dessus de l'eau. Pourtant, sa performance ne semble pas très bonne selon les résultats de notre étude et même en faisant une optimisation des paramètres de l'algorithme à partir de notre jeu de données, aucune amélioration n'est observée ($R^2=0,47$ et $RMSE=0,65$ pour l'ensemble du jeu de données et $R^2=0,64$ et $RMSE=0,58$ pour $\text{Log}([PC])$ estimés est au-dessus de 1,5). Le nouvel algorithme calibré génère en fait des résultats moins bons si on se fit à la majorité des tests statistiques présentés au Tableau 4 et au Tableau 5 (amélioration uniquement pour le coefficient de détermination pour le sous-ensemble de données où $\text{Log}([PC])$ est au-dessus de 1,5). Cette observation particulière peut s'expliquer par l'emploi du logarithme des valeurs des différents paramètres à l'étude. En fait, l'algorithme de Shalles et Yacobi a un nombre considérable d'estimations négatives (88 avant la calibration sur 298 données au total) qui sont éliminées du jeu de données lorsqu'une opération logarithme est appliquée et ce nombre diminue après la calibration (54 estimations de $[PC]$ négatives). Nous croyons que l'ajout de ces nouvelles estimations qui étaient négatives avant la calibration vient abaisser la performance de l'algorithme. Par contre, la diminution du nombre d'estimations négatives démontre qu'il y a déjà une amélioration de l'algorithme calibré malgré les mauvais résultats au niveau statistique.

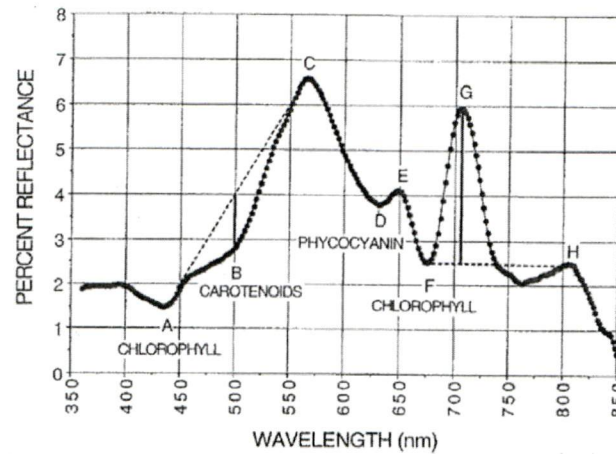


Figure 25 : Moyenne de 34 spectres de réflectance pris au lac Carter, Nebraska (tirée de Schalles et Yacobi (2000))

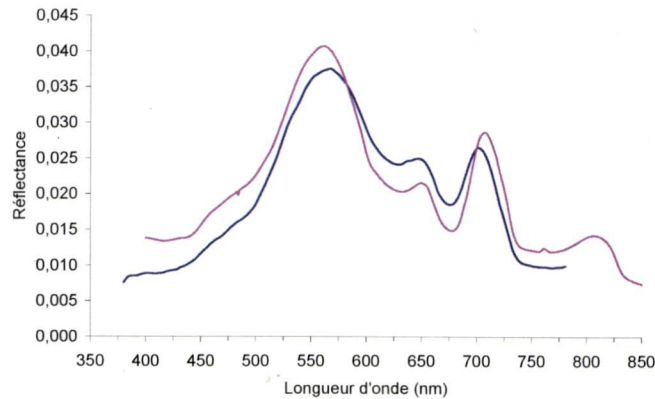


Figure 26 : Moyenne des spectres de réflectance des Pays-Bas (bleu) et de l'Espagne (violet) provenant de notre jeu de données fourni par Simis *et al.* (2006).

L'algorithme de rapport de bandes développé par Simis *et al.* (2005) donne de meilleurs résultats comparativement à ceux de Dekker (1993) selon le Tableau 4 et le Tableau 5. Cependant, toujours selon les statistiques recueillies, l'algorithme de Schalles et Yacobi (2000) a une meilleure précision (plus petit RMSE) lorsque toutes les données sont considérées comparativement à l'algorithme de Simis *et al.* (2005). Puis, pour le sous-ensemble de données au-dessus de 1,5, leurs précisions sont très semblables. Cette observation peut s'expliquer en grande partie par un effet limitatif de l'utilisation du logarithme des variables dans la situation présente : les estimations nulles ou négatives des [PC] estimées sont dans l'impossibilité de recevoir une transformation logarithmique et elles ont dû être mises de côté lors des analyses. Bien souvent, ces estimations négatives des algorithmes possèdent des erreurs relatives importantes et elles sous-estiment nécessairement la vraie valeur en [PC] qui ne peut être en réalité en dessous de zéro. En fait, l'algorithme de Schalles et Yacobi a un nombre important

d'estimations négatives qui n'ont pas pu être considérées dans notre analyse. Nous croyons que l'élimination de ces données apporte un certain biais en améliorant le coefficient de détermination et le RMSE de l'algorithme de rapport de bandes de Schalles et Yacobi. Des 298 données disponibles pour évaluer les algorithmes, 30% ont été éliminées pour l'algorithme de Schalles et Yacobi (2000), 15% pour celui Dekker (1993) et 4% pour Simis *et al.* (2005). Ainsi, seulement 4% des estimations de l'algorithme de Simis sont en dessous de zéro ce qui suggère une meilleure performance comparativement à l'algorithme de Schalles et Yacobi qui possède 30% d'estimations en [PC] négatives (valeurs qui sont impossibles en réalité). La pertinence des rapports de bandes employés dans l'algorithme de Simis *et al.* (2005) peut expliquer en partie les bons résultats obtenus. Par contre, il faut nuancer ce résultat, car le fait que cet algorithme semi-empirique fut développé à partir des données employées dans cette étude facilite aussi sa bonne performance, tout comme celle de notre algorithme de 2^e ordre. Notons aussi que l'algorithme de 2^e ordre développé dans cette recherche ne pouvait pas générer d'estimations négatives en raison de sa nature propre. Un autre résultat intéressant en lien avec l'algorithme de Simis *et al.* (2005) est l'amélioration obtenue à partir d'une simplification de leur algorithme semi-analytique (une seule relation empirique en fonction du rapport R708/R620 : voir la section 2.6.2 ci-dessus).

Enfin, il est possible de conclure grâce aux tests statistiques que notre algorithme de variabilité de 2^e ordre obtient des résultats équivalents ou meilleurs ($R^2=0,68$ et $RMSE=0,25$) par rapport aux algorithmes empiriques et l'algorithme semi-empirique de Simis *et al.* (2005). De plus, selon les statistiques du Tableau 6 pour les Log([PC]) estimés inférieurs à 1,5 ([PC] < 32 mg m⁻³), l'algorithme de 2^e ordre génère les meilleures estimations pour les faibles valeurs en [PC] qui sont très problématique chez les autres algorithmes. En plus des assez bons résultats générés, l'algorithme à l'étude est facile à développer. En effet, la méthodologie à suivre s'implémente sans trop de difficulté et le concept des anomalies, de la variabilité de 1^{er} ordre et de 2^e ordre se comprennent avec assez d'aisance.

Afin d'améliorer davantage cet algorithme, nous avons opté pour deux avenues. D'une part, nous avons raffiné l'algorithme de 2^e ordre en employant uniquement les valeurs de Log([PC]) estimés au-dessus de 1,5 dès la première étape. Ce seuil a maximisé l'optimisation des paramètres du nouvel algorithme pour ce sous-ensemble de données jugé au-dessus de la limite de détection

spectrale de la PC. Il génère les meilleurs résultats pour le sous-ensemble de données où $\text{Log}([PC])$ estimés est au-dessus de 1,5 ($R^2=0,72$ et $\text{RMSE}=0,19$). Le deuxième algorithme alternatif a été modifié au niveau des relations de variabilité de 1^{er} et de 2^e ordre. Ces étapes clés pour l'estimation de la [PC] semblaient présenter deux tendances distinctes entre les données des Pays-Bas et celles de l'Espagne (voir Figure 13 par exemple). C'est ainsi que deux droites de régression ont été générées pour former un nouvel algorithme mieux adapté aux conditions des étendues d'eau provenant des deux pays à l'étude. Nos résultats démontrent qu'il est le meilleur pour estimer la [PC] lorsque toutes les données sont considérées ($R^2=0,78$ et $\text{RMSE}=0,36$).

Malgré leur bonne performance, les deux derniers algorithmes présentés possèdent des inconvénients notables. Tout d'abord, l'algorithme utilisant uniquement les données de $\text{Log}([PC])$ estimés au-dessus de 1,5 est limité à des prédictions au-dessus de 32 mg/m³ de PC, ce qui est une quantité assez considérable dépendamment des étendues d'eau et de la période saisonnière. Ainsi, plusieurs situations présentant une quantité faible ou relativement modérée de cyanobactéries dans la colonne d'eau seront tout simplement hors de portée de cet algorithme. Quant à l'algorithme qui emploie deux régressions linéaires séparées selon le pays pour décrire la relation de 2^e ordre, sa principale faiblesse vient du fait qu'il nécessite l'application de deux équations différentes (ou deux algorithmes) pour estimer les [PC]. Donc, il y a encore une fois une certaine perte d'universalité au niveau de cet algorithme. On ne peut pas savoir quel algorithme utiliser dans une nouvelle région. Il est possible d'imaginer pour un futur système de surveillance qu'il faudrait autant d'algorithmes qu'il y a de groupes phytoplanctoniques ou d'étendues d'eau optiquement différents dans la zone à l'étude, ce qui complique l'opérabilité du système. Néanmoins, les algorithmes de 2^e ordre développés semblent très prometteurs et génèrent les meilleurs résultats observés dans cette recherche.

L'objectif de cette recherche n'était pas de vérifier la faisabilité d'utiliser des capteurs satellitaires, mais plutôt de développer un nouvel algorithme de télédétection plus précis pour l'estimation de la [PC] indépendamment des bandes actuellement disponibles sur les capteurs satellitaires. Par contre, il est très intéressant d'évaluer son applicabilité, car l'utilisation éventuelle d'images satellites permettrait d'obtenir un portrait global et périodique de la situation sur les lacs de la région à l'étude. Plusieurs études indiquent le bien fondé des images MERIS

pour l'étude des cyanobactéries principalement en raison de sa bande à 620 nm qui est centrée sur le maximum d'absorption de la PC (Kutser *et al.*, 2006; Simis *et al.*, 2005; Simis *et al.*, 2006; Wynne *et al.*, 2008; Wheeler *et al.*, *in press*). L'algorithme de Simis *et al.* (2005) est évidemment bien adapté à MERIS puisqu'il a été développé en prenant en compte la configuration spectrale du capteur. L'algorithme de Schalles et Yacobi (2000), qui emploie le rapport R650/R625 (voir équation 12), ne pourrait pas substituer convenablement le maximum local de réflectance à 650 nm avec la bande 7 à 665 nm de MERIS (Tableau 2 à la section 2.2.2), car un creux causé par l'absorption de la Chl *a* est présent à 665 nm. Quant à l'algorithme de Dekker (1993), il possède le même problème en raison de la réflectance à 648 nm qu'il utilise, mais en plus il exploite la réflectance à 600 nm qui n'a pas de correspondant avec le capteur satellitaire MERIS (voir équation 13). En ce qui a trait à notre algorithme, nous faisons face au même problème avec la réflectance à 650 nm qui est employée pour PCI (R650/R620). De plus, le capteur MERIS n'a pas de bande précisément à 676 nm ($CI=R708/R676$) et c'est ainsi que sa bande 8 à 681 nm devrait alors être employée. Malgré ces difficultés, nous avons créé un nouvel algorithme de 2^e ordre qui a pour but de tester l'applicabilité de MERIS avec notre algorithme. Nous avons employé cette fois-ci les bandes de MERIS sur les données de Simis *et al.* (2006). Ainsi, dans ce nouvel algorithme, CI est égale $R708/R681$ et PCI est égal à $665/620$ pour être conforme au capteur (bandes 9, 8, 7 et 6 respectivement selon le Tableau 2). Les résultats finaux aux Figure 27a et Figure 27b sont assez prometteurs (RMSE=0,44 pour toute les données et RMSE=0,28 pour $\text{Log}([PC])$ estimés au-dessus de 1,5) et ils semblent meilleurs que ceux des trois algorithmes tirés de la littérature (voir Tableau 4 et Tableau 5). Enfin, nous avons appliqué cet algorithme à une image MERIS du 29 août 2010 couvrant la Baie Missisquoi (Figure 28). Il est possible d'observer certaines formes et tendances sur l'image qui pourraient indiquer l'emplacement de cyanobactéries. Cependant, puisque nous n'avons pas des mesures de [PC] (uniquement 10 mesures de [Chl *a*]), il n'est pas possible de valider la véracité de ce résultat.

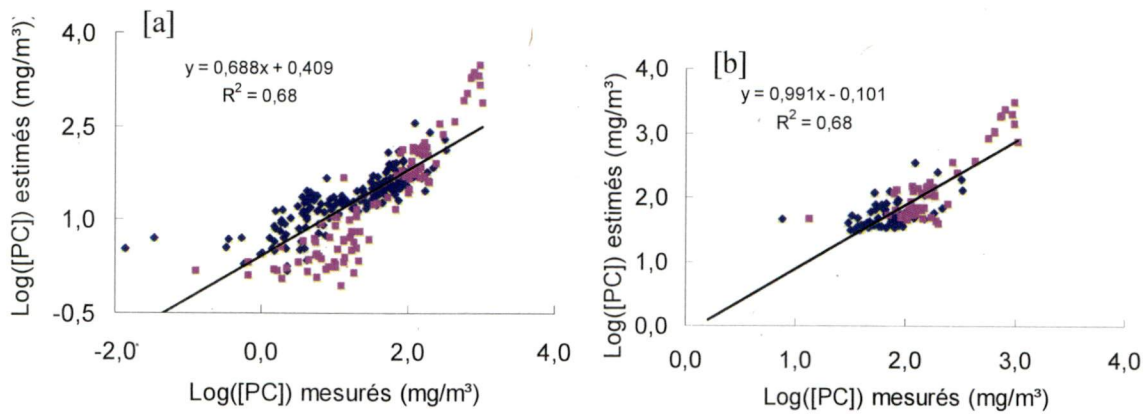


Figure 27 : Relation entre $\text{Log}([\text{PC}])$ estimés par notre algorithme de 2^e ordre développé à partir des bandes du capteur MERIS en fonction de $\text{Log}([\text{PC}])$ mesurés a) toutes les données et b) le sous-ensemble des données où $\text{Log}([\text{PC}])$ estimés est supérieur à 1,5. Résultats des Pays-Bas en bleu et ceux de l'Espagne en violet. RMSE=0,44 et 0,28 pour a) et b) respectivement.

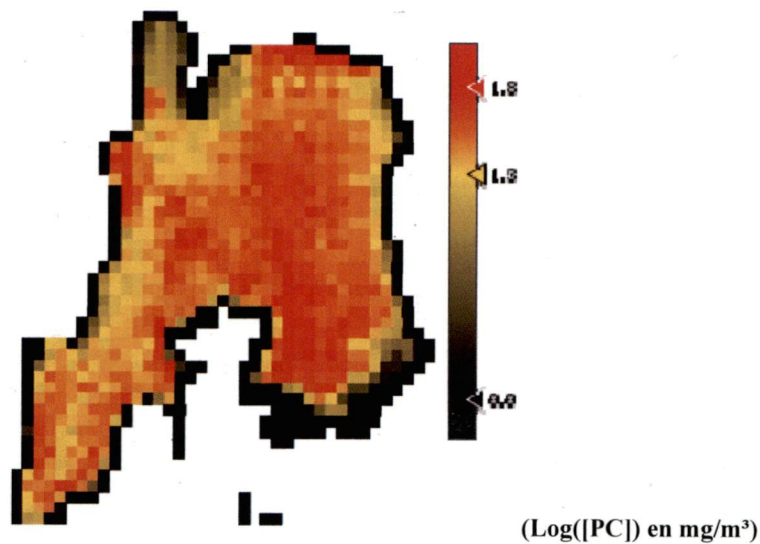


Figure 28 : Image MERIS du 29 août 2010 au-dessus de la baie Missisquoi présentant les estimations de $\text{Log}([\text{PC}])$ calculées à partir d'un algorithme de 2^e ordre développé à partir des bandes de MERIS et des données de Simis *et al.* (2006).

8. Conclusion

La méthodologie développée dans le cadre de cette recherche a permis d'améliorer la précision des estimations de la PC grâce à la nouvelle méthode de variabilité de 2^e ordre des spectres de réflectance de l'eau (précision mesurée en termes de RMSE et du coefficient de détermination par rapport à des concentrations en PC mesurées en laboratoire). Les trois autres algorithmes évalués provenant de la littérature ont présenté de plus grands écarts. Il a été démontré que l'ajout du terme de 2^e ordre améliore significativement notre algorithme en terme de R^2 . La méthode employée, qui consiste à soustraire la tendance principale d'un phénomène afin d'analyser plus facilement ses composantes secondaires (variabilités de 2^e ordre), a déjà fait ses preuves en milieu océanique (Alvain *et al.*, 2005; Brown *et al.*, 2008). Elle pourrait être appliquée à plusieurs autres problématiques en télédétection ou même à d'autres domaines de la science où l'on souhaiterait décomposer un phénomène global en différents éléments plus précis afin de les analyser plus facilement.

Par contre, l'algorithme de 2^e ordre comporte certaines faiblesses si les données de MERIS sont utilisées. Tout d'abord, puisque PCI (indice optique de la phycocyanine) emploie la réflectance à 650 nm (en plus de 620 nm), l'algorithme est difficilement applicable au capteur MERIS qui n'a pas la bande spectrale nécessaire. À ce niveau, l'algorithme de Simis *et al.* (2005) a un certain avantage, car il a été conçu selon les spécificités du capteur satellitaire. Une des façons de remédier à ce problème serait d'utiliser pour PCI le rapport R665/R620 compatible avec MERIS au lieu de R650/R620, ou bien d'explorer un nouveau rapport ou combinaison de bandes qui prend en compte les caractéristiques de MERIS, l'un des seuls capteurs satellitaires multispectraux ayant une bande centrée à 620 nm (maximum d'absorption de la PC). Ce dernier est présentement le meilleur capteur satellitaire disponible, mais il comporte tout de même certaines limitations au niveau de sa résolution spatiale qui est de 300 m. Par exemple, les concentrations en cyanobactéries peuvent varier très brusquement sur une même étendue d'eau (Reinart et Kutser, 2006; Kutser *et al.*, 2006) et des pixels de 300 m ne sont pas toujours adéquats pour observer cette variabilité spatiale (Kutser, 2004). De plus, au Québec, de nombreux lacs de petite taille sont touchés par les algues bleu-vert (Petit lac Lampton, lac Montjoie, Petit lac Saint-François, etc.). Pourtant, malgré certains inconvénients, la télédétection peut s'avérer une approche très avantageuse pour un plan de gestion des risques reliés aux cyanobactéries.

L'information trop fragmentaire qu'apportent les campagnes d'échantillonnages terrains n'est pas adéquate pour détecter le début, l'apogée et la fin d'une période de blooms de cyanobactéries, d'autant plus qu'elles sont incapables de donner un aperçu global du problème au niveau spatial. Les bouées et autres instruments prenant des mesures *in situ* en continu (bonne résolution temporelle) enregistrent toujours des mesures ponctuelles et on retrouve donc le même problème spatial. Cependant, avec la couverture spatiale et temporelle qu'offre la télédétection, cette problématique est résolue. D'ailleurs, cette constatation a été faite entre autres sur la baie Missisquoi durant une période de forts blooms (Wheeler *et al.*, *in press*).

L'une des limites majeures des algorithmes de télédétection de la PC que nous voulions améliorer est leur manque de précision pour les faibles et fortes valeurs en [PC]. Nous avons démontré qu'aucun des algorithmes testés (ceux de Dekker, 1993; Schalles et Yacobi, 2000; Simis *et al.*, 2005) permettent de réaliser des estimations précises pour des concentrations en PC en dessous de 32 mg/m^3 . Notre algorithme de 2^e ordre génère lui aussi de meilleurs résultats au-dessus de 32 mg/m^3 , mais il améliore la précision des estimations pour les faibles valeurs spécialement lorsque deux droites de régression de 2^e ordre sont employées pour les deux pays du jeu de données (Tableau 4). Par contre, cette solution n'est pas viable dans un contexte provincial ou national où plusieurs milieux lacustres différents pourraient être à l'étude. La recherche d'un algorithme plus universelle serait alors souhaitable. Selon WHO (2003), il y a un risque modéré sur la santé des êtres vivants lorsqu'environ 50 mg/m^3 de Chl *a* est retrouvée dans une eau dominée par les cyanobactéries. À ce stade, nous avons confiance que notre algorithme sera en mesure de prédire la présence et la quantité de cyanobactéries. Tout de même, il serait important d'avoir la capacité d'estimer la [PC], ou du moins avoir connaissance de leur présence, avant d'atteindre des concentrations de 50 mg/m^3 afin de prévenir les autorités et les personnes concernées par les risques potentiels. En effet, des risques faibles, mais potentiels sur la santé débutent à partir de 10 mg/m^3 (WHO, 2003). Ainsi, puisque l'étude par télédétection des faibles concentrations en PC semble avoir une moins bonne exactitude jusqu'à maintenant, d'autres études devraient se concentrer spécifiquement sur ce problème. Nous croyons que l'un des avancements important pour permettre davantage de recherche sur la télédétection des algues bleu-vert est l'exploration d'une méthode efficace et standard d'extraction du pigment PC. Sans la connaissance précise de la concentration de ce constituant optique déterminant, il est plus difficile de faire le lien direct

entre les cyanobactéries et le signal provenant des spectres de réflectance de l'eau. À notre connaissance, le jeu de données provenant de Simis *et al.* (2006) est l'un des rares à détenir un grand nombre de mesures de [PC] provenant de plusieurs lacs et réservoirs. Ainsi, un besoin se fait sentir à ce niveau. C'est d'ailleurs en raison des difficultés d'extraction de la PC que nous n'avons pas pu développer l'algorithme de variabilité de 2^e ordre avec des données acquises du Québec. Une fois les problèmes d'extraction résolus, il serait très intéressant d'appliquer notre nouvel algorithme de 2^e ordre à des spectres de réflectance pris sur les eaux de la baie Missisquoi par exemple pour valider sa précision et son potentiel. Il est évident que l'amélioration de l'exactitude des mesures de [PC] permettra de développer des algorithmes qui fourniront de meilleures prédictions pour les plus faibles valeurs de [PC]. Par contre, plusieurs autres facteurs importants sont en causes pour cette problématique. La qualité des mesures spectroradiométriques est primordiale aussi pour le développement et l'application d'algorithme de télédétection. Elle peut être affectée par plusieurs facteurs. Par exemple, le rapport signal sur bruit (rapport entre l'amplitude du signal de télédétection et l'amplitude du bruit ambiant ou des instruments) peut être faible pour les faibles valeurs en [PC]. De plus, la correction de l'effet de la réflexion spéculaire du ciel sur les données de réflectance l'eau n'est pas toujours adéquate et très exacte. Entre autres, en ce qui concerne les mesures acquises au Québec, il y a une surcorrection de l'apport du ciel ce qui entraîne des valeurs négatives de réflectances de l'eau pour les longueurs plus faibles du visible (voir Figure 31 de l'Annexe 2). De plus, malgré plusieurs précautions et un protocole strict suivi lors de la prise de mesure spectrale *in situ*, les réflexions spéculaires des rayons du soleil sur la surface de l'eau (reflet solaire ou "*sunglint*") ne peuvent pas toujours être évitées, et ceci modifie considérablement les spectres acquis. Tous ces facteurs influençant les mesures spectrales peuvent avoir des incidences importantes sur l'exactitude des estimations, en particulier pour les erreurs relatives des faibles concentrations. Un autre problème à considérer est la représentativité de l'échantillonnage de la [PC]. Les cyanobactéries ont la capacité de réguler leur flottabilité et elles se retrouvent souvent près de la surface ou à la surface de l'eau en période calme au lieu d'être mélangées uniformément dans la colonne d'eau (Kutser, 2004). Puis, lors des conditions de blooms, les navires de recherche ou certains instruments détruisent la distribution naturelle des cyanobactéries durant l'échantillonnage (Kutser, 2004). Il faut ainsi porter une attention particulière à ses caractéristiques propres à l'échantillonnage des algues bleu-vert. Enfin, il ne faut pas négliger

aussi les effets de certains paramètres optiques qui peuvent influencer et faire varier le spectre de réflectance de l'eau aux longueurs d'onde cruciales pour l'algorithme de 2^e ordre (dans le rouge et proche infrarouge). Par exemple, le CDOM, la rétrodiffusion et la matière en suspension sont tous susceptibles de modifier les mesures de réflectance près de 620 nm et ces variations du spectre seraient reliées à de fausses variations de [PC].

L'emploi de la télédétection pour la surveillance des cyanobactéries présente de nombreux avantages très intéressants et elle serait d'une grande utilité malgré le problème d'exactitude pour les faibles [PC]. L'un de ses avantages substantiels est qu'elle permet de faciliter l'étude de la distribution spatiale des blooms sur de grands territoires ce qui est beaucoup plus difficile, voire impossible, avec des mesures ponctuelles sur le terrain (Simis, 2006; Wynne *et al*, 2008). Sachant que la concentration au-dessus de 10 mg/m³ de PC est nocive pour la santé (WHO, 2003), un système de surveillance par télédétection qui aurait comme objectif premier de déceler la présence de la PC et d'alerter les autorités ne nécessiterait pas une grande exactitude pour l'estimation de la [PC]. Il aurait la possibilité d'être employé conjointement par un regroupement d'organismes gouvernementaux et de bassins versants. Une fois qu'un tel système est mis en place, son fonctionnement exige moins de personnes qualifiées qui sont habituellement indispensables pour échantillonner, traiter et analyser les échantillons récoltés sur le terrain. Par le fait même, le coût monétaire à long terme devrait être moins grand si la télédétection est employée. Ainsi, l'algorithme développé dans ce projet et son application par télédétection est un outil supplémentaire qui aide à donner un bilan plus complet de la situation de nos lacs tout en rassurant davantage la population qui est préoccupée par la présence et l'étendue des cyanobactéries depuis les dernières années au Québec.

9. Références

- Alvain, S., Moulin, C., Dandonneau, Y., et Breon, F. M. (2005) Remote sensing of phytoplankton groups in case 1 waters from global SeaWiFS imagery. *Deep-Sea Research Part I-Oceanographic Research Papers*, 52, 1989–2004.
- American Public Health Association (2005) Standard methods for the examination of water and wastewater. 21^e édition, les éditions Centennial, Baltimore, p. 1368
- ASD (2009) ASD Technical Guide. 3^e édition, Analytical Spectral Devices, Boulder (CO), 136p. Disponible à : www.asdi.com/tg_rev4_web.pdf
- Binding, C. E., Greenberg, T. A., Jerome, J. H., Bukata, R. P. et Letourneau, G. (2011) An assessment of MERIS algal products during an intense bloom in Lake of the Woods. *Journal of phytoplankton research*, vol. 33, n° 5, p. 793-806
- Bricaud, A., Morel, A. et Prieur, L. (1983) Optical efficiency factors of some phytoplankters. *Limnol Oceanogr* 28:816-832
- Bricaud, A., Claustre, H., Ras J. et Oubelkheir, K. (2004) Natural variability of phytoplankton absorption in oceanic waters: influence of the size structure of algal populations. *Journal of Geophys Res-Oceans* 109, **109**, C11010, doi:10.1029/2004JC002419
- Brown, C. A., Huot, Y., Werdell, P. J., Gentili, B. et Claustre, H., (2008) The origin and global distribution of second order variability in satellite ocean color and its potential applications to algorithm development. *Remote Sensing of Environment*, 112: 4186-4203
- Centre Saint-Laurent (2009) Les cyanobactéries dans le bassin Grands Lacs– Saint-Laurent *in* Infos Saint-Laurent, Environnement Canada, Gatineau, http://www.qc.ec.gc.ca/CSL/inf/inf073_f.html

Dekker, A. G. (1993) Detection of optical water quality parameters for eutrophic waters by high resolution remote sensing. Vrije Univ. Amsterdam

Doerffer, R. et Schiller, H. (2007). The MERIS Case 2 water algorithm. *International Journal of Remote Sensing*, Vol. 28, n° 3-4, pp. 517-535.

Environnement Canada (2011) La science et la technologie au service des Canadiens : Mettre au point de nouveaux outils pour surveiller la qualité des eaux intérieures du Canada, Ottawa, N° de catalogue : En39-1/2011F; ISBN 978-1-100-96414-0. 4 p.

Ernst, B., Dietz, L., Hoeger, S. J., Dietrich, D. R. (2005) Recovery of MC-LR in Fish Liver Tissue. *Environmental Toxicology*, vol. 20, no 4, p. 449-458.

Farabee, M., J. (2009) Chlorophyll and Accessory Pigments *in* Photosynthesis. Estrella Mountain Community College, Avondale,
<http://www.emc.maricopa.edu/faculty/farabee/BIOBK/BioBookPS.html#Chlorophyll>

Froidefond, J.M. et Doxaran, D. (2004) Télédétection optique appliquée à l'étude des eaux côtières, *Télédétection*, 2004, vol. 4, n° 2, p. 157–174

Gitelson, A., G. Garbuzov, F. Szilagyi, K-H. Mittenzwey, and A. Karnieli (1993) Quantitative Remote Sensing Methods for Real-time Monitoring Inland Water Quality, *Int. J. of Remote Sensing*, Vol. 14, pp. 1269-1295.

Gons, H. J. (1999) Optical teledetection of chlorophyll a in turbid inland waters. *Environmental Science and Technology*, n° 33, p. 1127–1132.

Gons, H. J., Rijkeboer, M., Bagheri, S., et Ruddick, K. G. (2000) Optical teledetection of chlorophyll a in estuarine and coastal waters. *Environmental Science and Technology*, n° 34, p. 5189–5192.

Gordon, H. R., Brown, O. B. et Jacobs, M., M. (1975) Computed relationships between the inherent and apparent optical properties of a flat homogeneous ocean. *Appl. Opt.* 14: 417–427

Gower, J. et King, S. (2007). Validation of chlorophyll fluorescence derived from MERIS on the west coast of Canada. *International Journal of Remote Sensing*, Vol. 28, Nos 3-4, pp. 625-635.

Gower, J. et King, S. (2008) New results from a global survey using MERIS MCI. *MERIS User Workshop*, Frascati, 22-26 Sept. 2008, 25 p.

<http://earth.esa.int/cgi-bin/confm8.pl?abstract=374>

ICES (2006) Report of the ICES-IOC Working Group on Harmful Algal Bloom Dynamics (WGHABD). ICES Oceanography committee, ICES; ICES WGHABD Report 2006, Pologne, 47 p., [www.ices.dk/reports/OCC/2006/WGHABD06.pdf].

Jeffrey S.W. et Vesk, M. (1997) Introduction to marine phytoplankton and their pigment signature. In Jeffrey SW, Mantoura RFC, Wright SW (eds.) *Phytoplankton pigments in oceanography*. UNESCO Publishing

Kutser, T. (2004) Quantitative detection of chlorophyll in cyanobacterial blooms by the satellite remote sensing. *Limnology and Oceanography*, 490 (6), 2179–2189.

Kutser, T., Metsamaa, L., Strömbeck, N. et Vahtmäe, E. (2006) Monitoring cyanobacterial blooms by satellite remote sensing. *Estuarine, Coastal and Shelf Science*, 67 : 303-312.

Lacasse Benoit, G. (2009) Le Lac Memphrémagog, un héritage à protéger!. *Memphrémagog Conservation, Magog*, <http://www.memphremagog.org/fr/index.php>

Lavoie, I., Laurion, I., Warren, A. et Vincent, W.F. (2007) Les fleurs d'eau de cyanobactéries, revue de littérature. *INRS rapport n. o. 916*, xiii, 124 p.

Matthews, MW, Bernard, S et Winter, K. (2010) Remote sensing of cyanobacteria-dominant algal blooms and water quality parameters in Zeekoevlei, a small hypertrophic lake, using MERIS. *Remote Sensing of Environment*, Vol. 114(9), pp 2070-2087

McKee, D., A. Cunningham et Dudek, A. (2007). Optical water type discrimination and tuning remote sensing band-ratio algorithms: Application to retrieval of chlorophyll and $K_d(490)$ in the Irish and Celtic Seas. *Estuarine, Coastal and Shelf Science*, 73, 827-834.

MDDEP – Ministère du Développement durable, Environnement et Parcs (2011) Bilan des lacs et cours d'eau touchés par une fleur d'eau d'algues bleu-vert au Québec, MDDEP, Québec, <http://www.mddep.gouv.qc.ca/eau/algues-bv/bilan/index.asp>

Mobley, C., D. (1994) *Light and Water: Radiative Transfer in Natural Waters*. Academic Press, San Diego, 592 p.

Moore, G. F., Aiken, J., et Lavender, S. J. (1999) The atmospheric correction of water colour and the quantitative retrieval of suspended particulate matter in Case II waters: application to MERIS. *Int. J. Remote Sens.*, 20, 1713-1733.

Morel, A. et Bricaud, A. (1981) Theoretical results concerning light absorption in a discrete medium, and application to specific absorption of phytoplankton. *Deep-Sea Res* 28:1375-1393

Morel, A. (1988) Optical modeling of the upper ocean in relation to its biogenous matter content (case 1 water). *Journal of Geophysical Research*, vol. 93, n° C9, p. 10749-10768

Morel, A. et Gentili, B. (2009) A simple band rapport technique to quantify the colored dissolved and detrital organic material from ocean color remotely sensed data. *Remote Sensing of Environment*, 113, 998-1011

Morel, A. et Prieur, L. (1977) Analysis of variations in ocean color. *Limnol. Oceanogr.*, 22(4), 709-722.

Mueller, J., L., Davis, C., Arnone, R., Frouin K., R., Carder, K, Lee, Z., P., Steward, R.G., Hooker, S., Mobley C., D, McLean, S. (2003) Above-Water Radiance and Remote Sensing Reflectance Measurement and Analysis Protocols *in* Mueller, J., L., Fargion, G., S. and McClain, C., R. (2003) Ocean Optics Protocols For Satellite Ocean Color Sensor Validation, Revision 4, Volume III: Radiometric Measurements and Data Analysis Protocol, National Aeronautical and Space administration, Greenbelt, 84 p.

NASA Science (2009) Ocean Color. NASA, Washington, <http://nasascience.nasa.gov/earth-science/oceanography/living-ocean/ocean-color>

Neville R. A., Gower, J. F. R. (1977), Passive remote sensing of phytoplankton via chlorophyll α fluorescence, *Jour. Geophys. Res.*, Vol. 82, No. 24, 3487-3493.

OBVBM - Organisme de bassin versant de la baie Missisquoi (2001) Le territoire du bassin versant de la baie Missisquoi, OBVBM, Bedford, <http://www.obvbm.org/territoire>

Park, Y-J. et Ruddick, K. (2007). Detecting algae blooms in european waters. *Envisat Symposium 2007 Proceeding*, Montreux, Switzerland, 23-27 April 2007.

Photo Research (2010) PR-650 SpectraScan[®] Colorimeter. Photo Research Inc, Chatsworth, <http://www.photoresearch.com/current/pr650.asp?type=0>

Reinart, A. et Kutser, T. (2006) Comparison of different satellite sensors in detecting cyanobacterial bloom events in the Baltic Sea. *Remote Sensing of Environment*, 102 : 74-85.

Ruiz-Verdú, A., Simis, S.G.H., de Hoyos, C., Gons, H.J. et Peña-Martínez, R. (2008). An evaluation of algorithms for the remote sensing of cyanobacterial biomass. *Remote Sensing of Environment*, vol. 112, n° 11, p. 3996-4008

Ryer, A. (1997) *Light Measurement Handbook*. International light, Newburyport, 60 p.

Schalles, J., F. et Yacobi, Y., Z. (2000) Remote detection and seasonal patterns of phycocyanin, carotenoid, and chlorophyll pigments in eutrophic waters. *Archives fur Hydrobiologia - Special Issues Advancements in Limnology*, 55: 153-168.

Siegel, M. R. (1961) *Statistics, Schaum's outline series in Mathematics*, McGraw-Hill

Simis, S., Peters, S.W.F. et Gons, J.H. (2005) Remote sensing of the cyanobacterial pigment phycocyanin in turbid inland water. *Limnology and Oceanography*, 50 : 237-242

Simis, S., A. Ruiz-Verdu, J. A. Dominguez-Gomez, R. Pena-Martinez, S. W. M. Peters et Gons, H. J.. (2006). Influence of phytoplankton pigment composition on remote sensing of cyanobacterial biomass. *Remote Sensing of Environment*, 106, 414-427.

Vincent, R. K., Qin, X., McKay, M. L., Miner, J., Czajkowskik., Savino, J. et Bridgeman, T. (2004) Phycocyanin detection from LANDSAT TM data for mapping cyanobacterial blooms in Lake Erie. *Remote Sensing of Environment*, 89 : 381-392.

Wheeler, S. M., Morrissey, L. A., Levine, S. N., Livingston, G. P., et Vincent, W. F. (in press) Mapping cyanobacterial blooms in Lake Champlain's Missisquoi Bay using QuickBird and MERIS satellite data. *Journal of Great Lakes Research*

WHO *pour* World Health Organization (2003) Guidelines for safe recreational water environments. Volume 1 : Coastal and Fresh Waters, WHO Library Cataloguing-in-Publication Data, Geneva, 219 p.

Wynne, T. T., Stumpf, R. P., Tomlinson, M. C., Warner, R. A., Tester, P. A., Dyble, J. et Fahnenstiel, G. L. (2008) Relating spectral shape to cyanobacterial blooms in the Laurentian Great Lakes. *International Journal of Remote Sensing*, 29:12, 3665 — 3672

Annexe 1 – Algorithme semi-analytique développé par Simis et al (2005)

$$1) a_{chl}(665) = (\{[R(709)/R(665)] \times [a_w(709) + b_b]\} - b_b - a_w(665)) \times \gamma^{-1}$$

Cette expression provient d'une simple application, aux longueurs d'onde à 665 et 709 nm, de la relation universelle de Gordon $R(0, \lambda) = f b_b / (a + b_b)$ où la variation spectrale de f et b_b est présumée négligeable par rapport au coefficient d'absorption.

$$2) a_{chl}(620) + a_{pc}(620) = (\{[R(709)/R(620)] \times [a_w(709) + b_b]\} - b_b - a_w(620)) \times \phi^{-1}$$

$$3) a_{pc}(620) = (\{[R(709)/R(620)] \times [a_w(709) + b_b]\} - b_b - a_w(620)) \times \phi^{-1} - [\epsilon \times a_{chl}(665)]$$

$$4) [PC] = a(620)_{pc} / a^*_{pc}(620)$$

Le Tableau 7 ci-dessous, de Simis *et al* (2005), donne la signification des symboles utilisés dans l'algorithme. Notons aussi que les symboles ϕ^{-1} , ϵ et γ^{-1} désignent :

γ^{-1} = facteur de correction qui est le rapport entre le coefficient d'absorption retrouvée à 665 nm par l'algorithme et celui mesuré pour les pigments à la longueur d'onde de 665 nm (pour la Chl *a*).

ϕ^{-1} = facteur de correction qui est le rapport entre le coefficient d'absorption retrouvée à 620 nm par l'algorithme et celui mesuré pour les pigments à la longueur d'onde de 620 nm (pour la PC).

ϵ = facteur qui permet d'estimer l'absorption de la Chl *a* à 620 nm à partir de son absorption à 665nm.

Table 1. List of symbols and abbreviations.

Symbol/abbreviation	Description	Units
$a_i(\lambda)$	Absorption coefficient of compound <i>i</i> at waveband λ . Subscripts used: <i>w</i> = water; chl = chlorophyll <i>a</i> ; PC = phycocyanin; tsm = total suspended matter; tripton = pigment-bleached TSM	m^{-1}
$a_i^*(\lambda)$	Specific absorption coefficient of pigment <i>i</i> at waveband λ	$m^2 (mg \text{ pigment})^{-1}$
b_b	Backscattering coefficient	m^{-1}
$R(0, \lambda)$	Subsurface irradiance reflectance at wavelength λ and depth 0	dimensionless
Chl <i>a</i>	Chlorophyll <i>a</i> (concentration)	$mg \text{ m}^{-3}$
PC	Phycocyanin (concentration)	$mg \text{ m}^{-3}$
CDOM	Colored dissolved organic matter	

Tableau 7 : Liste des symboles et abréviations (Simis *et al.*, 2005)

Annexe 2 – Correction de la réflexion de surface de l'éclairement du ciel

Grâce à la signature spectrale de la luminance du ciel normalisée par la luminance du Spectralon 5%, il est possible de minimiser l'effet de la réflexion spéculaire du ciel sur les données de réflectance de l'eau en presumant que (i) la réflectance à 900 nm est largement spéculaire et (ii) que la réflectance spéculaire est approximativement indépendante de la longueur d'onde. Ces présomptions nous permettent d'appliquer la correction suivante :

$$R_{\text{corr}}(\lambda) = R_{\text{eau}}(\lambda) - R_{\text{eau}}(900) * f_{\text{ciel}}(\lambda) / f_{\text{ciel}}(900) \quad (14)$$

où R_{corr} est la réflectance de l'eau corrigée, R_{eau} est la réflectance de l'eau et f_{ciel} est une estimation de la fraction de l'éclairement du ciel par rapport à l'éclairement total et λ représente les longueurs d'onde. Pour une démonstration complète de l'équation 14, voir l'annexe 3. Quatre figures sont présentées ci-dessous, soit la réflectance de l'eau [$R_{\text{eau}}(\lambda)$], la correction appliquée au spectre de l'eau (le terme $R_{\text{eau}}(900) * f_{\text{ciel}}(\lambda) / f_{\text{ciel}}(900)$), la réflectance de l'eau corrigée de la réflexion spéculaire du ciel ($R_{\text{corr}}(\lambda)$) et enfin une figure qui présente les trois spectres cités précédemment (Figure 29, Figure 30, Figure 31 et Figure 32 respectivement). Elles permettent d'illustrer les effets de la correction du ciel et elles proviennent de spectres acquis le 25 août 2009 à la station 2 de la baie Missisquoi.

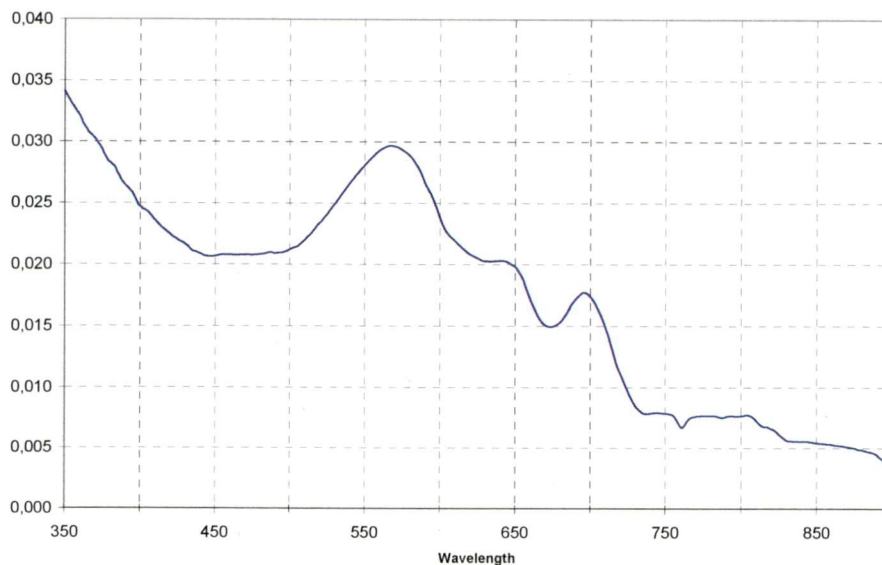


Figure 29 : Exemple de spectre de la réflectance de l'eau ($R_{\text{eau}}(\lambda)$)

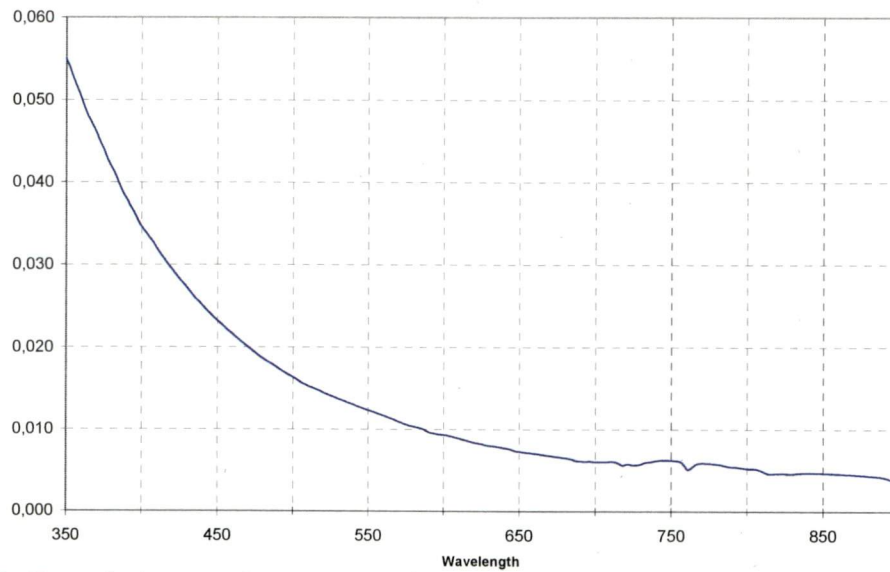


Figure 30 : Exemple de correction appliquée au spectre de la réflectance de l'eau ($R_{\text{eau}}(\lambda)$), soit R_{surf} de l'équation 18 et 19 ci-dessous à l'Annexe 3.

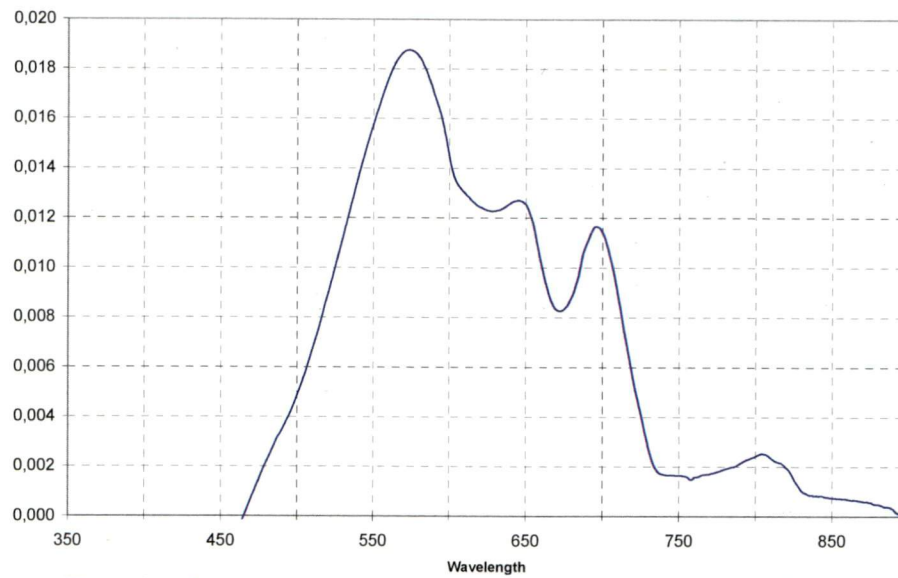


Figure 31 : Exemple de spectre de la réflectance de l'eau corrigée ($R_{\text{corr}}(\lambda)$)

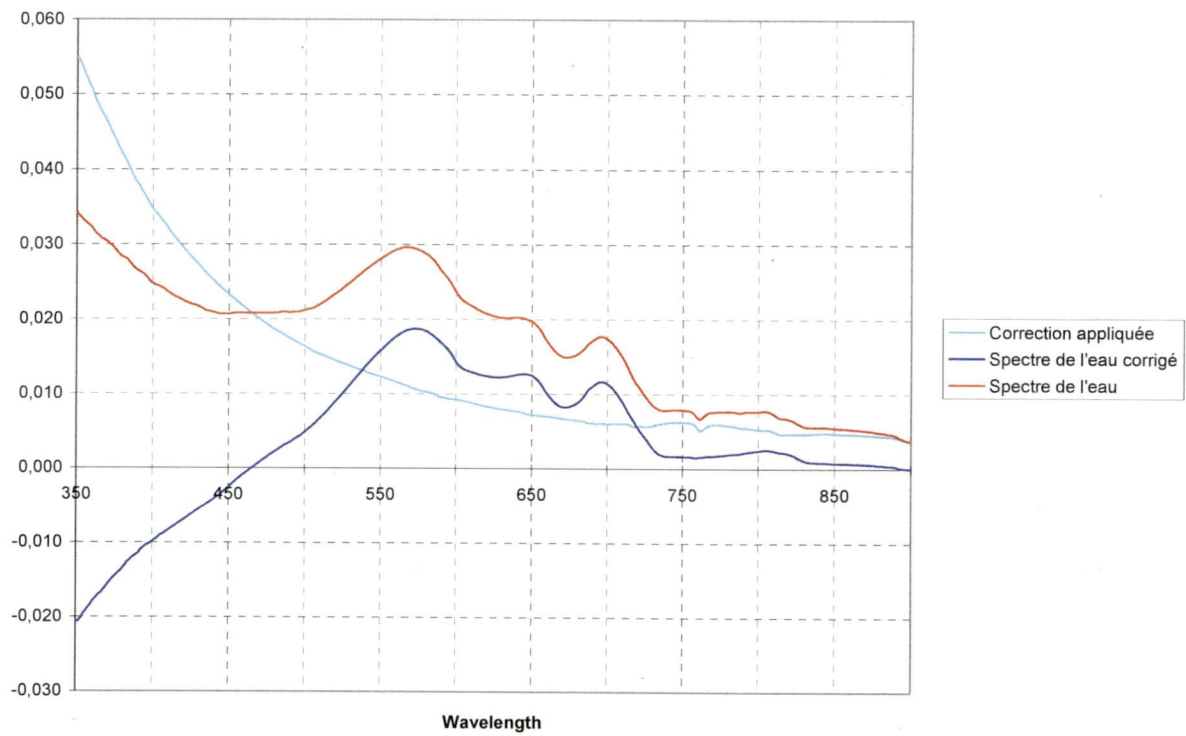


Figure 32 : Spectre de l'eau (rouge), spectre de l'eau corrigé (bleu foncé) et correction appliquée (bleu clair) pour un exemple concret de la baie Missisquoi (mesures prises le 25 août 2009 à la station 2).

Annexe 3 – Démonstration de la formule de correction du ciel

En assumant que le scintillement du soleil est négligeable à l'angle d'observation, la correction de la réflexion de surface (Fresnel) devrait être :

$$L_{\text{eau}} = L_{\text{meas}} - L_{\uparrow \text{ciel}} \quad (15a)$$

$$L_{\text{eau}} = L_{\text{meas}} - \rho_{\text{Fr}} L_{\text{ciel}} \quad (15b)$$

où ρ_{Fr} est la réflectance de Fresnel (pour un exemple voir l'équation (64) de Mueller et Austin (1995), malgré qu'il y a une erreur de signe dans l'équation). La Figure 33 suivante illustre les composantes du rayonnement en rapport avec le problème de réflexion du ciel à la surface de l'eau.

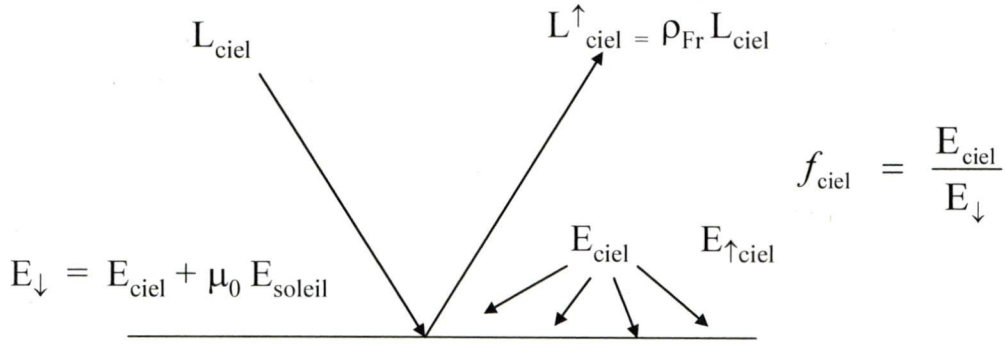


Figure 33 : Composantes du rayonnement concernant le problème de la réflexion de la luminance du ciel. La fraction de l'éclairement diffuse du ciel par rapport à l'éclairement total est représenté par f_{ciel} .

Pour les conditions du ciel sans nuage ou ennuagées de façon homogène, il est possible de faire l'approximation suivant : $E_{\text{ciel}} \cong \pi L_{\text{ciel}}$ où L_{ciel} est considérée comme une luminance du ciel dans une direction donnée. Il s'en suit que la luminance du ciel ascendante (réfléchie) peut s'écrire comme ceci $E_{\uparrow \text{ciel}} \cong \pi L_{\uparrow \text{ciel}}$. La moyenne (éclairement) du coefficient de Fresnel peut être écrite ainsi :

$$\bar{\rho}_{\text{Fr}} = \frac{E_{\uparrow \text{ciel}}}{E_{\text{ciel}}} \cong \frac{L_{\uparrow \text{ciel}}}{L_{\text{ciel}}} = \rho_{\text{Fr}} \quad (16)$$

Il est à noter que la réflectance de l'éclairement du ciel calculée à la surface est relative à l'éclairement descendant total :

$$R_{\text{surf}} = \frac{E_{\uparrow \text{ciel}}}{E_{\downarrow}} \cong \frac{\rho_{\text{Fr}} E_{\text{ciel}}}{E_{\downarrow}} \cong \rho_{\text{Fr}} f_{\text{ciel}} \quad (17)$$

où R_{surf} représente un estima de la réflectance du ciel dans la direction de prise de mesure (cette mesure représente exactement ce que l'on désire éliminer afin d'avoir une estimation des contributions dues aux interactions sous la surface uniquement). En multipliant l'équation (15b) par π / E_{\downarrow} nous obtenons effectivement :

$$R_{\text{eau}} = R_{\text{meas}} - \frac{\rho_{\text{Fr}} \pi L_{\text{ciel}}}{E_{\downarrow}} \cong R_{\text{meas}} - \frac{\rho_{\text{Fr}} E_{\text{ciel}}}{E_{\downarrow}} \cong R_{\text{meas}} - \rho_{\text{Fr}} f_{\text{ciel}} = R_{\text{meas}} - R_{\text{surf}} \quad (18)$$

Puis, si ρ_{Fr} est considérée comme étant approximativement indépendant de la longueur d'onde, alors la réflectance de surface à λ relativement à la réflectance de surface à λ_0 est donnée par :

$$R_{\text{surf}} = R_{\text{surf},0} \frac{f_{\text{ciel}}}{f_{\text{ciel},0}} \quad (19a)$$

$$\cong R_{\text{surf},0} \left(\frac{L_{\text{ciel}}}{L_{\text{ciel},0}} \right) \left(\frac{E_{\downarrow,0}}{E_{\downarrow}} \right) \quad (19b)$$

où on a utilisé l'équation 17 en plus de la propriété quasi isotropique (quasi Lambertienne) de la luminance du ciel. À partir des équations 18 et 19a, nous pouvons générer l'équation 20 ci-dessous ou l'équivalent de l'équation 14 de l'annexe 1 où λ_0 est une longueur d'onde de référence dans le proche infrarouge. La figure conceptuelle suivante (Figure 34) démontre que le rapport de la fraction de l'éclairement du ciel (le 3^e terme de l'équation 19b) va généralement générer une contribution de R_{surf} plus faible comparativement au rapport de luminance (le 2^e terme de l'équation 19b).

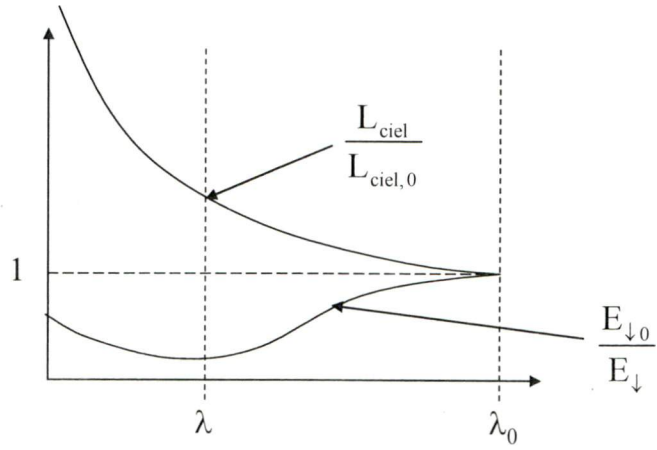


Figure 34 : Représentation schématique des rapports de l'équation 19b.

Finalement, pour une certaine réflectance à une longueur d'onde du proche infrarouge (R_0), on peut assumer que cette réflectance est dominée par l'effet de surface selon l'équation suivante :

$$R_{\text{eau}} = R_{\text{meas}} - R_{\text{surf},0} \frac{f_{\text{ciel}}}{f_{\text{ciel},0}} \quad (20)$$

Annexe 4 - Analyses spectrales supplémentaires sur les anomalies

À première vue, l'algorithme de variabilité de 2^e ordre donne des résultats prometteurs pour les concentrations en PC > 32 mg/m³. Par contre, il est important d'étudier en détail si les anomalies (résidus) de la Figure 15 observées autour de la droite moyenne entre Log([PC]) et Log(CI) peuvent être corrélées davantage avec les mesures de réflectance. Cet exercice permet d'explorer d'autres avenues pour l'analyse spectroradiométrique des anomalies de la [PC] (Δ [PC] ou Δ Log([PC]) dans notre cas) qui peuvent, selon notre hypothèse, améliorer les algorithmes de télédétection des cyanobactéries.

De façon concrète, nous allons examiner dans cette section si les anomalies de la PC (Δ Log([PC])) varient significativement avec des réflectances ou des rapports de réflectances dans le visible, autres que les anomalies de PCI (R650/R620) étudiées dans la section précédente. Dans le cas contraire, il serait possible d'affirmer que les variations autour de la tendance moyenne n'ont pas le potentiel d'être un meilleur indicateur spectral lorsqu'elles sont jumelées à d'autres bandes spectrales du visible.

Les figures 35 à 39 montrent la relation entre Δ Log([PC]) et R488/R556, R700/R560, R700/R668, R650/R620 et R708/R620 respectivement. Les rapports R488/R556 (O'Reilly *et al.*, 1998), R700/R560 (Giterson *et al.*, 1993) et R700/R668 (Matthews *et al.*, 2010) peuvent potentiellement être corrélés avec Δ Log([PC]) puisqu'ils ont tous été employés précédemment pour estimer la biomasse algale. De plus, les rapports R650/R620 (Schalles et Yacobi, 2000) et R708/R620 (Simis *et al.*, 2005) proviennent d'algorithme estimant la concentration en [PC]. Par contre, il est facile de constater qu'il n'y a pas de dépendance entre les anomalies de Log([PC]) et les cinq rapports de bandes précédents. Pour les figures 40 à 43, les Δ Log([PC]) sont maintenant comparés à de simples bandes de réflectance, soit R568, R620, R664 et R776 respectivement. Encore une fois, aucune corrélation significative ne semble exister entre Δ Log([PC]) et les mesures spectrales. Ainsi, ces derniers résultats démontrent que les anomalies obtenues suite à la soustraction de la variabilité de 1^{er} ordre ne peuvent pas s'expliquer par nos mesures de réflectance obtenues de façon *in situ*, et ce, pour notre ensemble de données.

Puisque nous n'avons pas trouvé de relation significative entre $\Delta\text{Log}([\text{PC}])$ et les bandes et rapports de bandes de réflectance étudiés, nous pouvons conclure que l'indice de chlorophylle ($\text{CI}=\text{R}708/\text{R}676$) utilisé pour obtenir la variabilité de 1^{er} ordre est présentement le meilleur indicateur de la présence de la PC dans notre jeu de données (Figure 13, $R^2=0,66$). Ainsi, avec un simple rapport de bandes qui tente d'évaluer la $[\text{Chl } a]$, il est possible d'estimer la variabilité de 1^{er} ordre de la $[\text{PC}]$ avec une précision relativement bonne. Cependant, nous savons que la $[\text{PC}]$ n'est pas toujours corrélée linéairement avec la $[\text{Chl } a]$ dans les lacs.

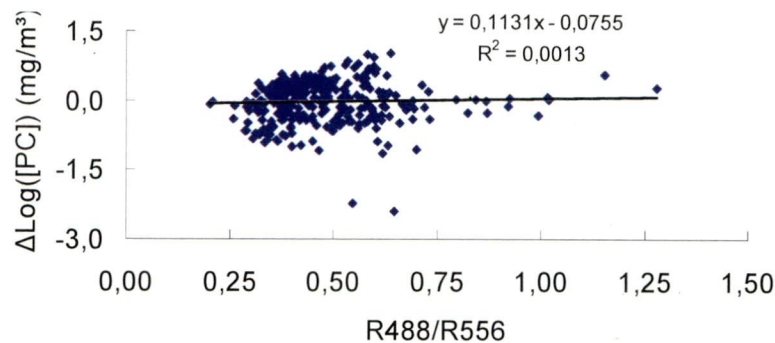


Figure 35 : Relation entre $\Delta\text{Log}([\text{PC}])$ et $\text{R}488/\text{R}556$ pour les données des Pays-Bas et de l'Espagne.

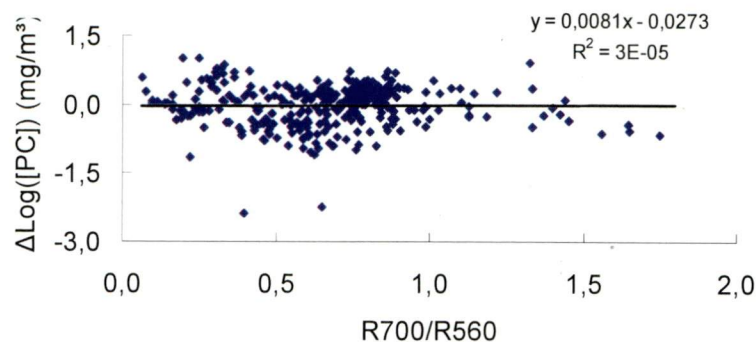


Figure 36 : Relation entre $\Delta\text{Log}([\text{PC}])$ et $\text{R}700/\text{R}560$ pour les données des Pays-Bas et de l'Espagne.

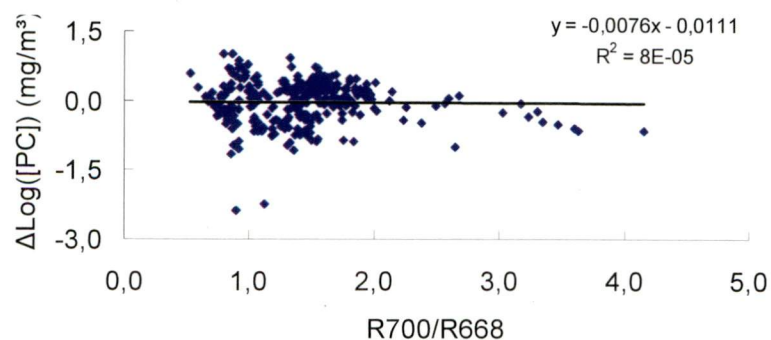


Figure 37 : Relation entre $\Delta\text{Log}([\text{PC}])$ et $\text{R}700/\text{R}668$ pour les données des Pays-Bas et de l'Espagne.

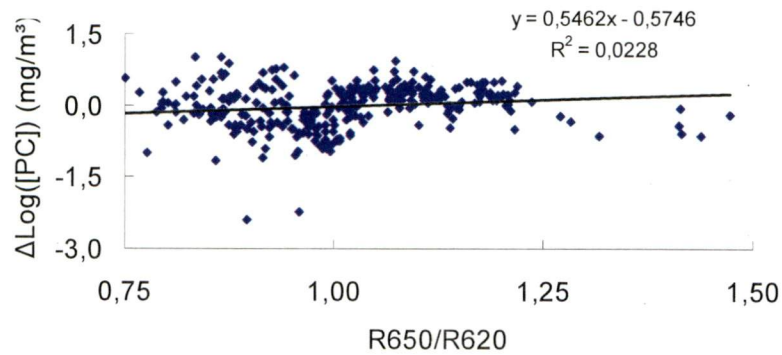


Figure 38 : Relation entre $\Delta\text{Log}([\text{PC}])$ et $R650/R620$ pour les données des Pays-Bas et de l'Espagne.

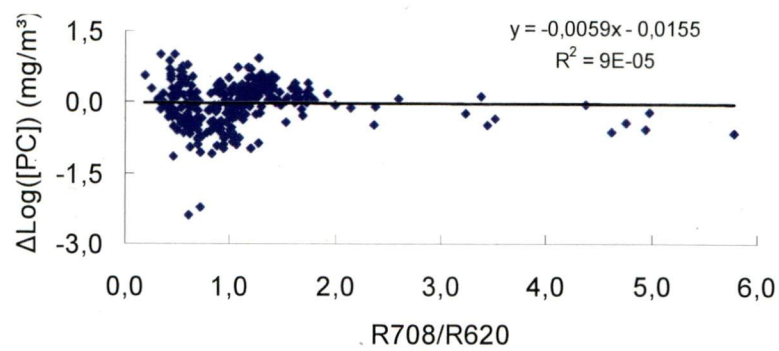


Figure 39 : Relation entre $\Delta\text{Log}([\text{PC}])$ et $R708/R620$ pour les données des Pays-Bas et de l'Espagne.

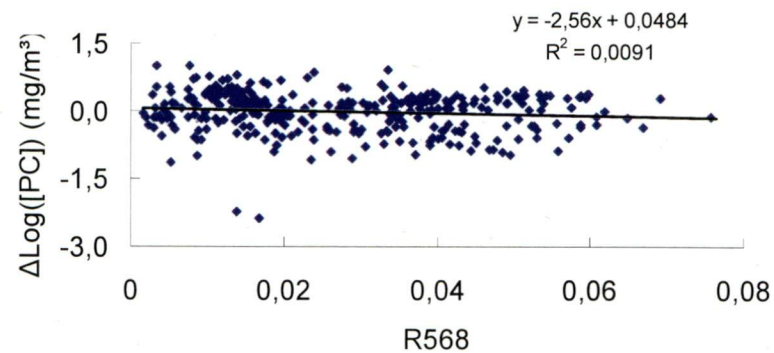


Figure 40 : Relation entre $\Delta\text{Log}([\text{PC}])$ et $R568$ pour les données des Pays-Bas et de l'Espagne.

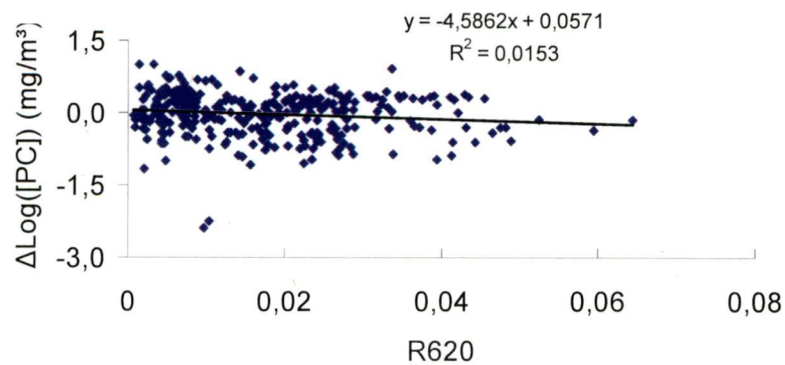


Figure 41 : Relation entre $\Delta\text{Log}([\text{PC}])$ et R620 pour les données des Pays-Bas et de l'Espagne.

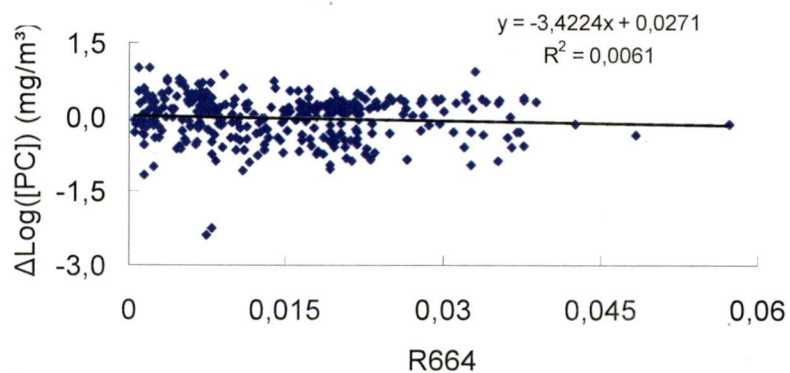


Figure 42 : Relation entre $\Delta\text{Log}([\text{PC}])$ et R664 pour les données des Pays-Bas et de l'Espagne.

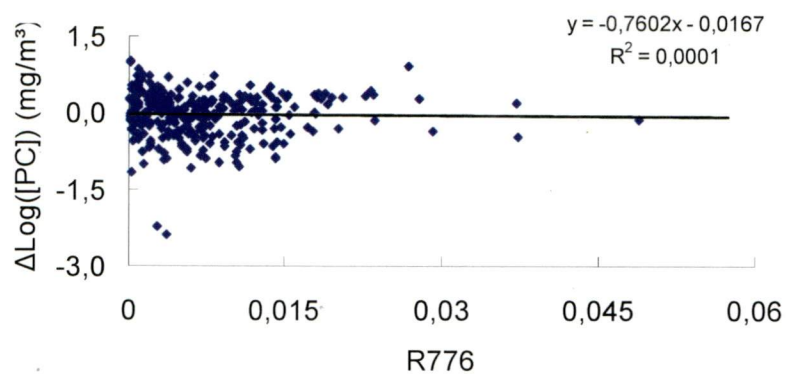


Figure 43 : Relation entre $\Delta\text{Log}([\text{PC}])$ et R776 pour les données des Pays-Bas et de l'Espagne.

Annexe 5 – Deux algorithmes de 2e ordre alternatifs

Deux autres avenues sont considérées pour améliorer notre algorithme de 2^e ordre initial. Tout d'abord, il a été observé que lorsque les valeurs de $\text{Log}([\text{PC}])$ estimés sont supérieures à 1,5, notre algorithme semble générer des résultats plus significatifs. C'est ainsi qu'un nouvel algorithme est créé en utilisant la même méthodologie qu'auparavant, mais en employant uniquement les données de $\text{Log}([\text{PC}])$ estimés au-dessus de 1,5 dès le début de son développement (les valeurs estimées de $\text{Log}([\text{PC}])$ au-dessus de 1,5 ont été déterminées suite aux résultats de la Figure 16 du premier algorithme de 2^e ordre). La Figure 44 et la Figure 45 illustrent les deux relations de 1^{er} ordre, la Figure 46 représente la relation de 2^e ordre de l'algorithme et enfin, la Figure 47 montre le résultat final. Pour ces résultats, les deux données représentées en rouge sur les figures ont été éliminées du jeu de données puisqu'elles ont été jugées comme étant des observations aberrantes. Les estimations de $\text{Log}([\text{PC}])$ se sont améliorées en terme de RMSE (RMSE=0,19) si on les compare à celles illustrées à la Figure 16b (RMSE=0,25) générées par l'algorithme développé à partir de toutes les données (les valeurs de R^2 n'ont pas changées significativement). Notons ici que le terme de 2^e ordre ($\Delta[\text{PC}]$) n'apporte pas d'amélioration significative pour cet algorithme (pour l'algorithme de 1^{er} ordre, le RMSE=0,19 et $R^2=0,72$).

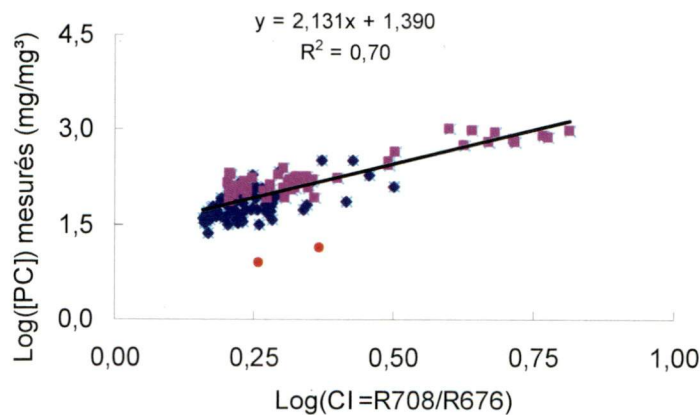


Figure 44 : Relation de 1^{er} ordre entre $\text{Log}([\text{PC}])$ et $\text{Log}(\text{CI})$ pour le sous-ensemble des données où $\text{Log}([\text{PC}])$ estimés est supérieur à 1,5. Résultats des Pays-Bas en bleu et ceux de l'Espagne en violet. Les deux points rouges sont des données aberrantes et elles ont été éliminées de l'analyse.

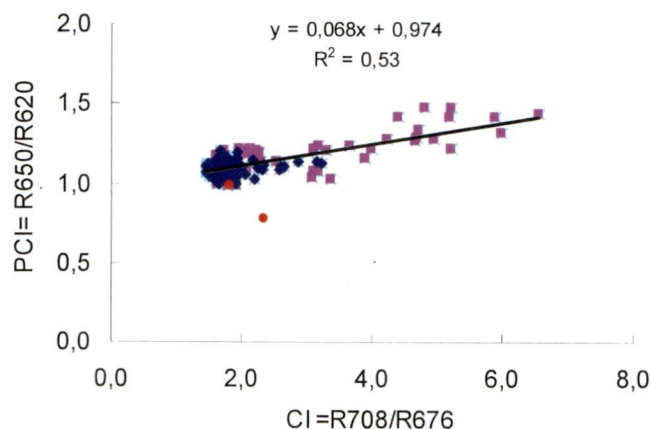


Figure 45 : Relation de 1^{er} ordre entre PCI et CI pour le sous-ensemble des données où Log([PC]) estimés est supérieur à 1,5. Résultats des Pays-Bas en bleu et ceux de l'Espagne en violet. Les deux points rouges sont des données aberrantes et elles ont été éliminées de l'analyse.

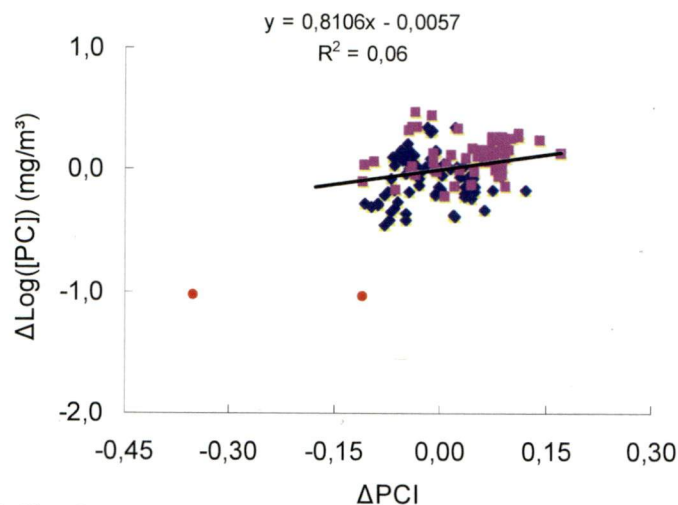


Figure 46 : Relation de 2^e ordre entre ΔLog([PC]) et ΔPCI pour le sous-ensemble des données où Log([PC]) estimés est supérieur à 1,5. Résultats des Pays-Bas en bleu et ceux de l'Espagne en violet. Les deux points rouges sont des données aberrantes et elles ont été éliminées de l'analyse.

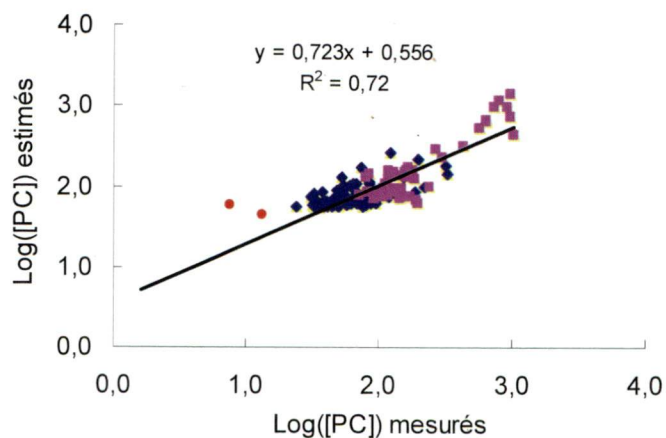


Figure 47 : Relation entre Log([PC]) estimés par notre algorithme (équation 5), développé à partir du sous-ensemble des données où Log([PC]) estimés est supérieur à 1,5, en fonction de Log([PC]) mesurés. Résultats des Pays-Bas en bleu et ceux de l'Espagne en violet. Les deux points rouges sont des données aberrantes et elles ont été éliminées de l'analyse. RMSE=0,19

Il semble y avoir deux tendances distinctes à la Figure 13 de notre algorithme entre les données des Pays-Bas et celles d'Espagne. Suite à cette observation, une nouvelle analyse consistant à évaluer deux relations différentes de 1^{er} et de 2^e ordre pour chaque site à l'étude (Pays-Bas et Espagne) est effectuée dans le but d'améliorer la performance de notre algorithme. La Figure 48, Figure 49 a et Figure 49 b montrent respectivement les deux relations de 1^{er} ordre entre les $\text{Log}([\text{PC}])$ et $\text{Log}(\text{CI})$ et les nouveaux résultats finaux engendrés pour toutes les données et pour celles dont $\text{Log}([\text{PC}])$ estimées est supérieur à 1,5. Les deux autres étapes de la méthodologie ne sont pas présentées ci-dessous puisqu'elles sont très semblables à la Figure 14 et Figure 15 mise à part (deux relations de 1^{er} ordre entre PCI et CI et deux relations de 2^e ordre entre les anomalies pour chacun des sites à l'étude). Ce dernier algorithme est le plus performant lorsque l'ensemble du jeu de données est considéré ($R^2=0,78$ et $\text{RMSE}=0,36$), car il semble augmenter la précision des estimations pour les plus faibles valeurs de $\text{Log}([\text{PC}])$ comparativement aux algorithmes de 2^e ordre et ceux de la littérature [significativement différent selon le test de Siegel (1961) lorsque l'ensemble du jeu de données est considéré].

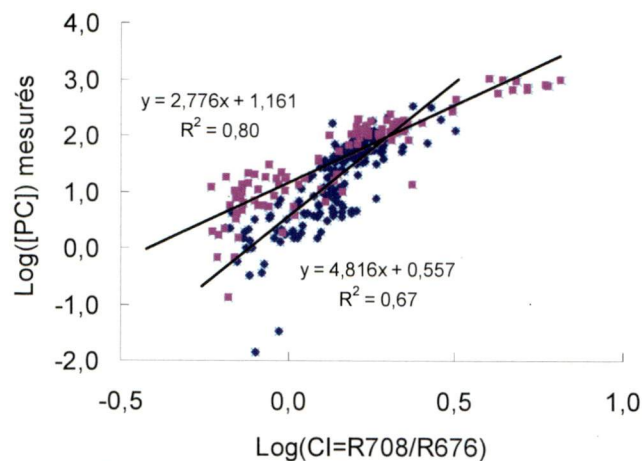


Figure 48 : Deux relations de 1^{er} ordre entre $\text{Log}([\text{PC}])$ et $\text{Log}(\text{CI})$, l'une pour les données des Pays-Bas en bleu et l'autre pour les données d'Espagne en violet.

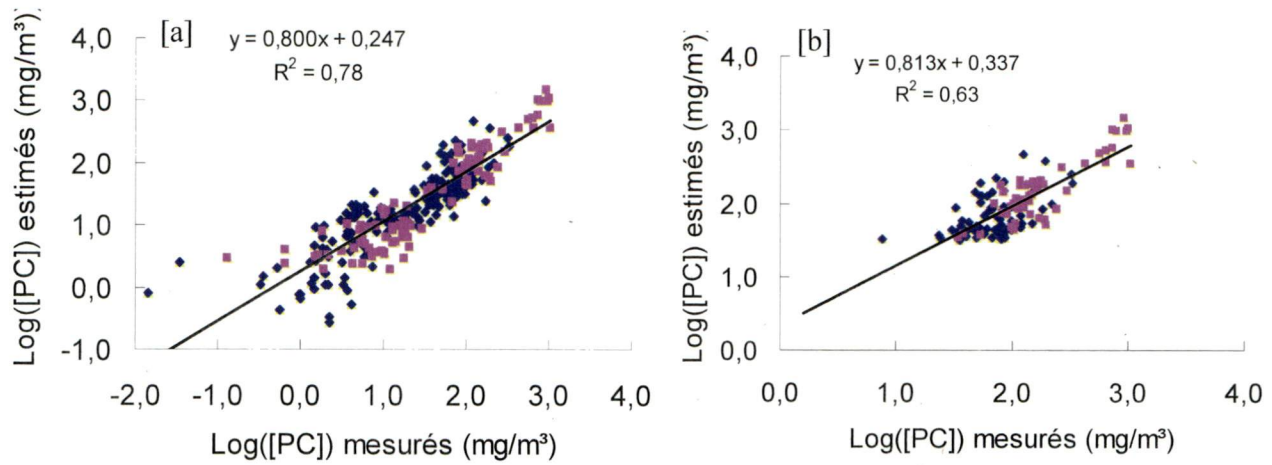


Figure 49 : Relation entre Log([PC]) estimés par notre algorithme (équation 5) en fonction de Log([PC]) mesurés lorsque deux relations de 2^e ordre sont considérées pour a) toutes les données et b) le sous-ensemble des données où Log([PC]) estimés est supérieur à 1,5. Résultats des Pays-Bas en bleu et ceux de l'Espagne en violet. RMSE=0,36 et 0,24 pour a) et b) respectivement.

Annexe 6 – Relation entre les erreurs relatives et le rapport des concentrations en pigments

Une analyse entre les erreurs relatives des estimations [$\text{Log}([\text{PC}])$ estimés - $\text{Log}([\text{PC}])$ mesurés] et le rapport des concentrations mesurées des pigments [$\text{Log}([\text{PC}] / [\text{Chl } a])$ pour montrer la proportion entre les deux pigments] vient renchérir les observations au niveau de la performance des algorithmes. La relation entre ces deux paramètres est présentée aux figures 50 à 53 respectivement pour les algorithmes de Dekker (1993), Schalles et Yacobi (2000), Simis *et al.* (2005) et notre algorithme de 2^e ordre [a] pour la totalité du jeu de données et b) pour le sous-ensemble au-dessus de 1,5]. Il est possible d'observer une relation négative entre les deux paramètres pour chacun des algorithmes lorsque toutes les données sont considérées. Ce résultat indique d'une part que tous les algorithmes surestiment de façon importante la concentration en PC lorsque le rapport $\text{Log}([\text{PC}] / [\text{Chl } a])$ est faible. Ce phénomène peut s'expliquer par le fait que les algorithmes sont mal adaptés aux situations où il y a une absence du creux d'absorption à 620 nm (ou un faible creux d'absorption) lorsqu'il n'y a pas ou presque pas de PC. Ainsi, en absence ou quasi-absence de PC, les estimations deviennent fortement corrélées avec la $[\text{Chl } a]$ et il n'y a pas de facteur correctif qui considère la faible proportion du pigment réellement ciblé, soit la PC. La relation entre $\text{Log}([\text{PC}])$ estimés inférieurs à 1,5 et la $[\text{Chl } a]$ présenté à la Figure 54 et Figure 55 (algorithme de 2^e ordre et celui de Simis respectivement) démontre ces derniers propos. Bien que l'algorithme de Simis *et al.* (2005) ai une faible corrélation avec la $[\text{Chl } a]$ mesurée pour les faibles valeurs en $[\text{PC}]$ estimée, elle est d'autant plus faible lorsque la $[\text{PC}]$ mesurée est analysée (voir Figure 24 et Tableau 6 de la section Résultat). Par ailleurs, ce manque de sensibilité par rapport à la $[\text{PC}]$ se reflète beaucoup moins pour notre algorithme de 2^e ordre lorsque l'on élimine les $\text{Log}([\text{PC}])$ estimés en dessous de 1,5, tel qu'il est observable à la Figure 53b (faible corrélation et coefficient de détermination). Il est possible de croire que le terme de la variabilité de 2^e ordre est adéquat et bien adapté à l'étude des variations de la $[\text{PC}]$ lorsque celle-ci est au-dessus de 32 mg/m³. Quant aux trois autres algorithmes de la littérature, ils semblent toujours fortement influencés par la $[\text{Chl } a]$ et seraient donc plus sensible aux variations de la $[\text{Chl } a]$ que celles de la $[\text{PC}]$. Dans un deuxième temps, il est possible de conclure selon ces derniers résultats que les quatre algorithmes semblent générer des erreurs plus faibles lorsque le rapport $\text{Log}([\text{PC}]/[\text{Chl } a])$ est élevé, soit lorsque les cyanobactéries deviennent de plus en plus dominantes à la surface de l'eau. Cette observation confirme d'autres résultats observés plus tôt dans ce document et dans la recherche de Simis *et al.* (2006). Notons que le nombre de

données analysées pour chacun de ces résultats varie en fonction du nombre d'estimations négatives des algorithmes de la littérature (logarithme d'une valeur négatif n'est pas défini dans les nombres réelles) et en fonction des $\text{Log}([\text{PC}])$ estimés au-dessus de 1,5 qui diffèrent d'un algorithme à l'autre.

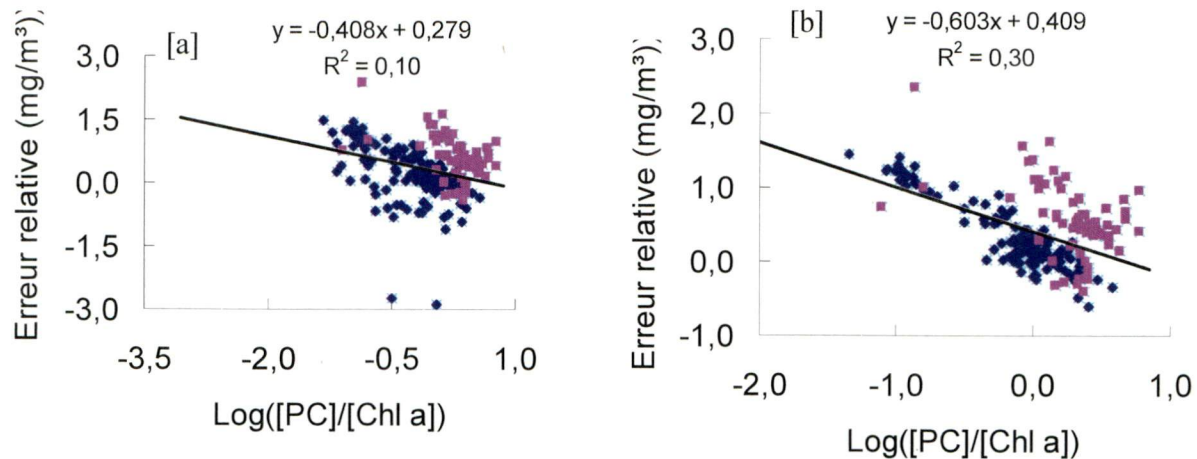


Figure 50 : Relation entre les erreurs relatives des estimations de $\text{Log}([\text{PC}])$ de l'algorithme de Dekker (1993) en fonction du rapport des concentrations mesurées des pigments, soit $\text{Log}([\text{PC}] / [\text{Chl } a])$, pour a) toutes les données et b) le sous-ensemble des données où $\text{Log}([\text{PC}])$ estimés est supérieur à 1,5. Résultats des Pays-Bas en bleu et ceux de l'Espagne en violet.

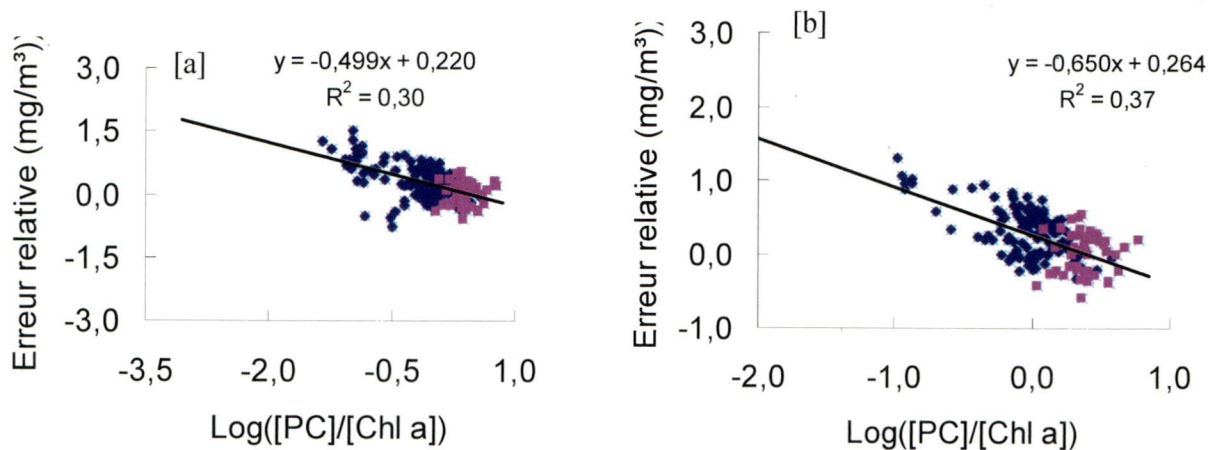


Figure 51 : Relation entre les erreurs relatives des estimations de $\text{Log}([\text{PC}])$ de l'algorithme de Schalles et Yacobi (2000) en fonction du rapport des concentrations mesurées des pigments, soit $\text{Log}([\text{PC}] / [\text{Chl } a])$, pour a) toutes les données et b) le sous-ensemble des données où $\text{Log}([\text{PC}])$ estimés est supérieur à 1,5. Résultats des Pays-Bas en bleu et ceux de l'Espagne en violet.

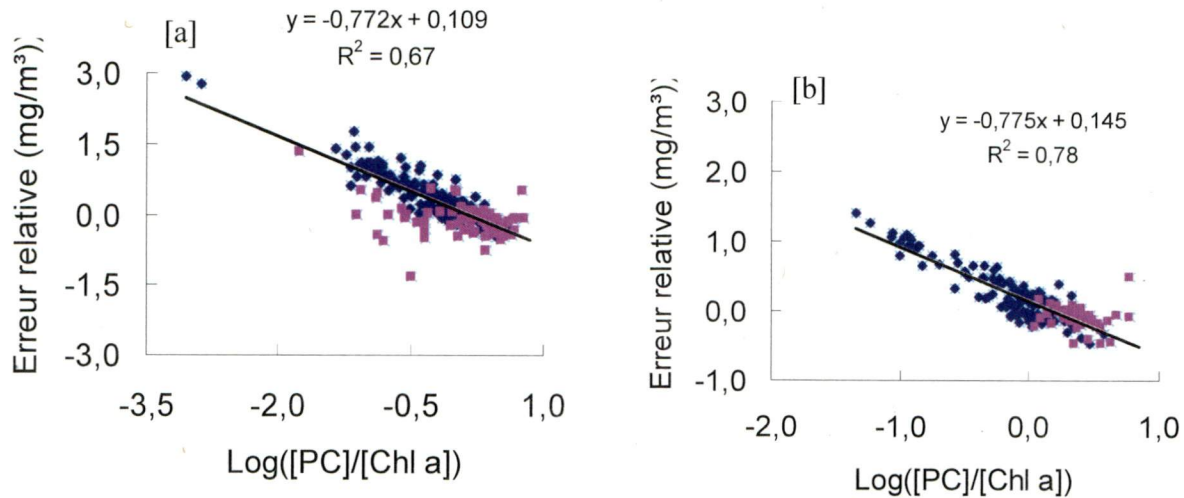


Figure 52 : Relation entre les erreurs relatives des estimations de Log([PC]) de l'algorithme de Simis *et al.* (2005) en fonction du rapport des concentrations mesurées des pigments, soit Log([PC] / [Chl a]), pour a) toutes les données et b) le sous-ensemble des données où Log([PC]) estimés est supérieur à 1,5. Résultats des Pays-Bas en bleu et ceux de l'Espagne en violet.

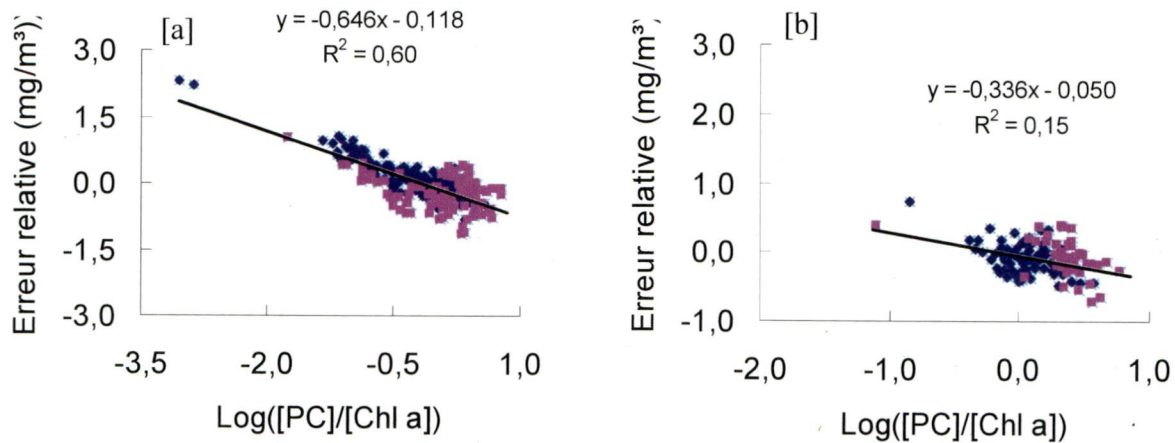


Figure 53 : Relation entre les erreurs relatives des estimations de Log([PC]) de notre algorithme de variabilité 2^e ordre en fonction du rapport des concentrations mesurées des pigments, soit Log([PC] / [Chl a]), pour a) toutes les données et b) le sous-ensemble des données où Log([PC]) estimés est supérieur à 1,5. Résultats des Pays-Bas en bleu et ceux de l'Espagne en violet.

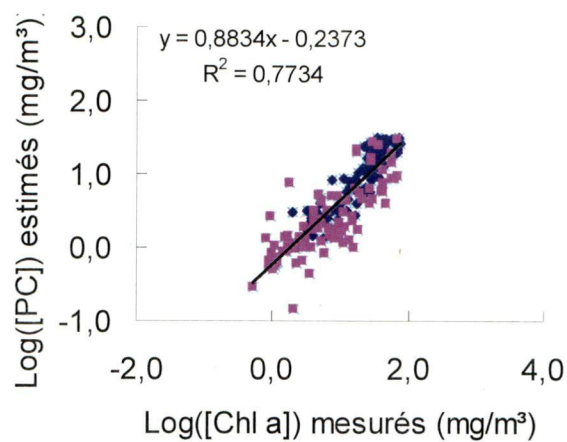


Figure 54 : Relation entre $\text{Log}([\text{PC}])$ estimés par l'algorithme de 2^e ordre et $\text{Log}([\text{Chl } a])$ mesurés pour le sous-ensemble des données où $\text{Log}([\text{PC}])$ estimés est inférieur à 1,5 ($[\text{PC}] < 32\text{mg}/\text{m}^3$). RMSE = 0,45.

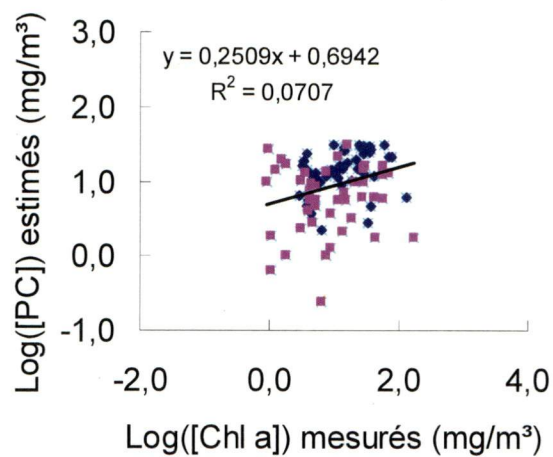


Figure 55 : Relation entre $\text{Log}([\text{PC}])$ estimés par l'algorithme de Simis *et al.* et $\text{Log}([\text{Chl } a])$ mesurés pour le sous-ensemble des données où $\text{Log}([\text{PC}])$ estimés est inférieur à 1,5 ($[\text{PC}] < 32\text{mg}/\text{m}^3$). RMSE = 0,59.